

OM protein - protein search, using sw model

Run on: March 5, 2004, 16:14:14 ; Search time 6 Seconds  
(without alignments)  
282.548 Million cell updates/sec

Title: US-10-057-890A-15  
Perfect score: 42  
Sequence: 1 GHHHHS 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query				ID	Description
		Match	Length	DB	%		
1	42	100.0	138	5	ABG32539	ABG32539	Human CCR
2	42	100.0	157	5	ABG32540	ABG32540	Human CCR
3	42	100.0	323	6	ABM67261	ABM67261	Photorhab
4	42	100.0	339	3	AAG44195	AAG44195	Arabidops
5	42	100.0	353	3	AAG44194	AAG44194	Arabidops
6	42	100.0	692	6	ABR40877	ABR40877	Oryza sat
7	42	100.0	1056	4	ABB70257	ABB70257	Drosophil
8	39	92.9	68	3	AAG59221	AAG59221	Arabidops
9	39	92.9	69	3	AAG59220	AAG59220	Arabidops

10	39	92.9	70	3	AAG26198	Aag26198	Arabidops
11	39	92.9	94	3	AAG59219	Aag59219	Arabidops
12	39	92.9	95	3	AAG26197	Aag26197	Arabidops
13	39	92.9	239	4	ABB70310	Abb70310	Drosophil
14	39	92.9	324	3	AAG39403	Aag39403	Arabidops
15	39	92.9	351	3	AAG39402	Aag39402	Arabidops
16	39	92.9	378	2	AAW38182	Aaw38182	Arabidops
17	39	92.9	378	3	AAB28599	Aab28599	Protein e
18	39	92.9	379	3	AAB28580	Aab28580	Arabidops
19	39	92.9	422	2	AAW71639	Aaw71639	Omega-cyc
20	39	92.9	605	3	AAG39401	Aag39401	Arabidops
21	39	92.9	1480	5	AAU77228	Aau77228	Human NR2
22	39	92.9	1484	2	AAR66040	Aar66040	Human N-m
23	39	92.9	1484	2	AAR80971	Aar80971	Human exc
24	39	92.9	1484	2	AAR92507	Aar92507	Human NMD
25	39	92.9	1484	2	AAW87510	Aaw87510	Human N-m
26	39	92.9	1484	3	AAV56134	Aay56134	Human NMD
27	39	92.9	1484	3	AAB26236	Aab26236	Human N-m
28	39	92.9	1484	5	ABB56529	Abb56529	Human NMD
29	39	92.9	1484	5	AAO18084	Aao18084	Human NMD
30	39	92.9	1484	6	ABU65899	Abu65899	Human N-m
31	39	92.9	1484	6	ABU61446	Abu61446	Human N-m
32	39	92.9	1484	6	ABU57681	Abu57681	Human N-m
33	39	92.9	1484	6	ABO43091	Abo43091	Human NMD
34	39	92.9	1484	7	AAE39257	Aae39257	Human NMD
35	39	92.9	1484	7	ADE62168	Ade62168	Human Pro
36	38	90.5	6	2	AAR23933	Aar23933	Lactoferr
37	38	90.5	6	6	ABP97637	Abp97637	Proteolyt
38	38	90.5	8	3	AAB00161	Aab00161	Hexahisti
39	38	90.5	9	2	AAW06142	Aaw06142	Variant a
40	38	90.5	9	2	AAV28306	Aay28306	Amino aci
41	38	90.5	9	3	AAV44463	Aay44463	Gly(His)8
42	38	90.5	9	4	AAB82692	Aab82692	His-tag.
43	38	90.5	9	4	AAE12682	Aae12682	Polyhisti
44	38	90.5	9	5	AAU84236	Aau84236	Molecular
45	38	90.5	9	6	ABU89706	Abu89706	Poly-hist

#### ALIGNMENTS

##### RESULT 1

ABG32539

ID ABG32539 standard; protein; 138 AA.

XX

AC ABG32539;

XX

DT 15-NOV-2002 (first entry)

XX

DE Human CCR5-based scaffolded fusion protein #1.

XX

KW Scaffolded protein; CCR5; HIV; human immunodeficiency virus infection;

KW ECD; extracellular domain; metal chelating motif; zinc finger protein;

KW integral membrane protein; soluble loop; intracellular domain; ICD;

KW gene therapy; immunogen; viral infection; human.

XX

OS Homo sapiens.

OS Synthetic.  
 XX  
 PN WO200260477-A1.  
 XX  
 PD 08-AUG-2002.  
 XX  
 PF 29-JAN-2002; 2002WO-US002377.  
 XX  
 PR 31-JAN-2001; 2001US-0265782P.  
 PR 31-JAN-2001; 2001US-0265858P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Coleman TA, Mansfield B;  
 XX  
 DR WPI; 2002-643357/69.  
 XX  
 PT Novel scaffolded fusion polypeptide useful for therapeutic purposes or  
 PT for screening molecules that bind/activate/inhibit/modulate the  
 PT polypeptide, comprises a functional polypeptide domain fused to a  
 PT scaffold domain.  
 XX  
 PS Example 1; Page 21; 64pp; English.  
 XX  
 CC The invention relates to a scaffolded fusion polypeptide comprising a  
 CC functional polypeptide domain fused to a scaffold domain, where the  
 CC functional polypeptide domain corresponds to a soluble loop of an  
 CC integral membrane protein (e.g. human CCR5, a transmembrane receptor  
 CC involved in HIV (human immunodeficiency virus) infection). Also included  
 CC are; (1) a polypeptide comprising a scaffold domain; (2) a nucleic acid  
 CC encoding the fusion polypeptide; (3) a vector cassette for the expression  
 CC of the fusion polypeptide comprising an expression region operably linked  
 CC to a promoter, where the expression region comprises a number of  
 CC cassettes, each of which encodes a module, domain or strand of the fusion  
 CC polypeptide and (4) a host cell comprising the vector or nucleic acid.  
 CC The fusion polypeptide is useful for screening molecules that  
 CC bind/activate/inhibit/modulate the fusion polypeptide, by expressing the  
 CC fusion polypeptide from and identifying a molecule that binds to the  
 CC fusion polypeptide. The fusion polypeptide is useful in diagnostic  
 CC methods, in assays to identify compounds that interact with loops of  
 CC fragments of an extracellular domain (ECD) or an intracellular domain  
 CC (ICD) or to rapidly assay the function of mutated portions of mutant  
 CC integral membrane proteins without having to produce significant  
 CC quantities of the entire mutant integral membrane protein, to generate  
 CC antibodies that recognise the integral membrane proteins from which they  
 CC are designed, to competitively bind the ligand of a naturally occurring  
 CC receptor in vitro or in vivo, to display and/or screen soluble domains  
 CC from protein such as integral membrane proteins, to probe the structure  
 CC of ECD or ICD, or both, of an integral protein membrane, to modulate the  
 CC activity of a receptor in vivo, and for treating or preventing viral  
 CC infection, preferably human HIV infection e.g. by gene therapy using the  
 CC encoding nucleic acid. The present sequence is a scaffolded protein based  
 CC on the ECD region of human CCR5 (not defined)  
 XX  
 SQ Sequence 138 AA;

Query Match 100.0%; Score 42; DB 5; Length 138;

Best Local Similarity 100.0%; Pred. No. 4.3;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
|||||  
Db 55 GHHHHS 60

RESULT 2

ABG32540

ID ABG32540 standard; protein; 157 AA.

XX

AC ABG32540;

XX

DT 15-NOV-2002 (first entry)

XX

DE Human CCR5-based scaffolded fusion protein #2.

XX

KW Scaffolded protein; CCR5; HIV; human immunodeficiency virus infection;  
KW ECD; extracellular domain; metal chelating motif; zinc finger protein;  
KW integral membrane protein; soluble loop; intracellular domain; ICD;  
KW gene therapy; immunogen; viral infection; human.

XX

OS Homo sapiens.

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Peptide 1. .19

FT /label= Signal\_peptide

FT Protein 20. .157

FT /label= Mature\_scaffolded\_protein

XX

PN WO200260477-A1.

XX

PD 08-AUG-2002.

XX

PF 29-JAN-2002; 2002WO-US002377.

XX

PR 31-JAN-2001; 2001US-0265782P.

PR 31-JAN-2001; 2001US-0265858P.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

XX

PI Coleman TA, Mansfield B;

XX

DR WPI; 2002-643357/69.

XX

PT Novel scaffolded fusion polypeptide useful for therapeutic purposes or  
PT for screening molecules that bind/activate/inhibit/modulate the  
PT polypeptide, comprises a functional polypeptide domain fused to a  
PT scaffold domain.

XX

PS Example 2; Page 41; 64pp; English.

XX

CC The invention relates to a scaffolded fusion polypeptide comprising a  
CC functional polypeptide domain fused to a scaffold domain, where the  
CC functional polypeptide domain corresponds to a soluble loop of an

CC integral membrane protein (e.g. human CCR5, a transmembrane receptor  
 CC involved in HIV (human immunodeficiency virus) infection). Also included  
 CC are; (1) a polypeptide comprising a scaffold domain; (2) a nucleic acid  
 CC encoding the fusion polypeptide; (3) a vector cassette for the expression  
 CC of the fusion polypeptide comprising an expression region operably linked  
 CC to a promoter, where the expression region comprises a number of  
 CC cassettes, each of which encodes a module, domain or strand of the fusion  
 CC polypeptide and (4) a host cell comprising the vector or nucleic acid.  
 CC The fusion polypeptide is useful for screening molecules that  
 CC bind/activate/inhibit/modulate the fusion polypeptide, by expressing the  
 CC fusion polypeptide from and identifying a molecule that binds to the  
 CC fusion polypeptide. The fusion polypeptide is useful in diagnostic  
 CC methods, in assays to identify compounds that interact with loops of  
 CC fragments of an extracellular domain (ECD) or an intracellular domain  
 CC (ICD) or to rapidly assay the function of mutated portions of mutant  
 CC integral membrane proteins without having to produce significant  
 CC quantities of the entire mutant integral membrane protein, to generate  
 CC antibodies that recognise the integral membrane proteins from which they  
 CC are designed, to competitively bind the ligand of a naturally occurring  
 CC receptor in vitro or in vivo, to display and/or screen soluble domains  
 CC from protein such as integral membrane proteins, to probe the structure  
 CC of ECD or ICD, or both, of an integral protein membrane, to modulate the  
 CC activity of a receptor in vivo, and for treating or preventing viral  
 CC infection, preferably human HIV infection e.g. by gene therapy using the  
 CC encoding nucleic acid. The present sequence is a scaffolded protein based  
 CC on the ECD region of human CCR5 (not defined)  
 XX  
 SQ Sequence 157 AA;

Query Match 100.0%; Score 42; DB 5; Length 157;  
 Best Local Similarity 100.0%; Pred. No. 4.9;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
 |||||  
 Db 74 GHHHHS 79

# RESULT 3

ABM67261

ID ABM67261 standard; protein; 323 AA.

XX

AC ABM67261;

XX

DT 20-NOV-2003 (first entry)

XX

DE Photorhabdus luminescens protein sequence #358.

XX

KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;  
 KW detection; food; gene expression; plant; animal; microorganism; toxin;  
 KW antibiotic; biopesticide; virulence factor; disease model; plague;  
 KW whooping cough.

XX

OS Photorhabdus luminescens.

XX

PN WO200294867-A2.

XX

PD 28-NOV-2002.  
 XX  
 PF 07-FEB-2002; 2002WO-IB003040.  
 XX  
 PR 07-FEB-2001; 2001FR-00001659.  
 XX  
 PA (INSP ) INST PASTEUR.  
 PA (CNRS ) CNRS CENT NAT RECH SCI.  
 XX  
 PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;  
 PI Buchrieser C;  
 XX  
 DR WPI; 2003-148459/14.  
 XX  
 PT Genomic sequence of Photorhabdus luminescens and encoded polypeptides,  
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.  
 XX  
 PS Claim 2; SEQ ID NO 358; 1205pp; French.  
 XX  
 CC The invention relates to the isolation of genes and their encoded  
 CC proteins from Photorhabdus luminescens. The isolated sequences are  
 CC sources of probes and primers for detecting the genome of P. luminescens  
 CC and related species; to study polymorphisms; for gene analysis and for  
 CC detection/amplification of the genes. Antibodies (Ab) raised against the  
 CC polypeptides encoded by the genes are used for detection/identification  
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that  
 CC carry a gene-containing vector are used to select compounds that  
 CC modulate, regulate, induce or inhibit expression of the genes in plants,  
 CC animals or microorganisms other than P. luminescens and are able to alter  
 CC response or sensitivity to toxins and antibiotics produced by P.  
 CC luminescens. Cells transformed to express the genes are useful for  
 CC recombinant production of the proteins, particularly toxins and  
 CC antibacterials useful as insecticides, bactericides and fungicides. The  
 CC genes, proteins, vectors containing the genes and Ab are also useful  
 CC therapeutically (to treat microbial infection by bacteria or fungi that  
 CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as  
 CC biopesticides. Other uses of the genes and the proteins are as virulence  
 CC factors and for identifying targets of human diseases for which P.  
 CC luminescens is a model (particularly plague and whooping cough). This  
 CC sequence represents one of the isolated P. luminescens proteins  
 XX  
 SQ Sequence 323 AA;

Query Match 100.0%; Score 42; DB 6; Length 323;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
 |||||  
 Db 191 GHHHHS 196

RESULT 4  
 AAG44195  
 ID AAG44195 standard; protein; 339 AA.  
 XX  
 AC AAG44195;

XX  
 DT 18-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 55329.  
 XX  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-00301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 23-APR-1999; 99US-0130891P.  
 PR 28-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 30-APR-1999; 99US-0132407P.  
 PR 04-MAY-1999; 99US-0132484P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 06-MAY-1999; 99US-0132487P.  
 PR 07-MAY-1999; 99US-0132863P.  
 PR 11-MAY-1999; 99US-0134256P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.  
 PR 04-JUN-1999; 99US-0137502P.  
 PR 07-JUN-1999; 99US-0137724P.  
 PR 08-JUN-1999; 99US-0138094P.

PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.



PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.

PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 42; DB 3; Length 339;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
 |||||  
 Db 164 GHHHHS 169

# RESULT 5

AAG44194

ID AAG44194 standard; protein; 353 AA.

XX

AC AAG44194;

XX

DT 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 55328.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-00301439.

XX

PR	25-FEB-1999;	99US-0121825P.
PR	05-MAR-1999;	99US-0123180P.
PR	09-MAR-1999;	99US-0123548P.
PR	23-MAR-1999;	99US-0125788P.
PR	25-MAR-1999;	99US-0126264P.
PR	29-MAR-1999;	99US-0126785P.
PR	01-APR-1999;	99US-0127462P.
PR	06-APR-1999;	99US-0128234P.
PR	08-APR-1999;	99US-0128714P.
PR	16-APR-1999;	99US-0129845P.
PR	19-APR-1999;	99US-0130077P.
PR	21-APR-1999;	99US-0130449P.
PR	23-APR-1999;	99US-0130510P.
PR	23-APR-1999;	99US-0130891P.
PR	28-APR-1999;	99US-0131449P.
PR	30-APR-1999;	99US-0132048P.
PR	30-APR-1999;	99US-0132407P.
PR	04-MAY-1999;	99US-0132484P.
PR	05-MAY-1999;	99US-0132485P.
PR	06-MAY-1999;	99US-0132486P.
PR	06-MAY-1999;	99US-0132487P.
PR	07-MAY-1999;	99US-0132863P.
PR	11-MAY-1999;	99US-0134256P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.

PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.

PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.

PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 100.0%; Score 42; DB 3; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
 |||||  
 Db 178 GHHHHS 183

# RESULT 6

ABR40877

ID ABR40877 standard; protein; 692 AA.

XX

AC ABR40877;

XX

DT 16-MAY-2003 (first entry)

XX

DE Oryza sativa oil trait related protein sequence SEQ ID NO:530.

XX

KW Plant; oil trait; oil phenotype; altered lipid profile; MAP kinase;  
 KW receptor-like protein kinase; mitogen activated protein kinase; oil;  
 KW LIP15-like transcription factor caleosin; ATP citrate lyase; SNF1;  
 KW CKC-like transcription factor; antisense inhibition; co-suppression;  
 KW transgenic plant.

XX

OS Oryza sativa.

XX

PN WO2003002751-A2.

XX

PD 09-JAN-2003.

XX

PF 27-JUN-2002; 2002WO-US020152.

XX

PR 29-JUN-2001; 2001US-0301913P.

XX

PA (DUPO ) DU PONT DE NEMOURS & CO E I.

PA (PION-) PIONEER HI-BRED INT INC.

XX

PI Allen SM, Allen WB, Cahoon RE, Epelbaum S, Famodu OO, Harvell LT;  
 PI Jones TJ, Kinney AJ, Klein TM, Li C, Oliveira IC, Sakai H, Shen B;  
 PI Tarczynski MC;

XX

DR WPI; 2003-201509/19.

XX

PT Novel nucleotide fragment encoding polypeptides having receptor-like  
 PT protein kinase activity, caleosin-like activity, useful for altering oil

PT phenotypes in plants such as sunflower, coconut, soybean, wheat and rice.  
 XX  
 PS Claim 12; Page 538-540; 542pp; English.  
 XX  
 CC The present invention describes an isolated nucleotide fragment (I)  
 CC comprising a nucleic acid sequence (NS) chosen from a NS encoding a  
 CC polypeptide (PP) having receptor-like protein kinase activity, mitogen  
 CC activated protein (MAP)-kinase activity, LIP15-like transcription factor  
 CC activity, caleosin-like activity, ATP citrate lyase activity, SNF1-like  
 CC activity and CKC-like transcription factor activity. Also described: (1)  
 CC complement (II) of (I); (2) a chimeric construct (III) comprising (I) or  
 CC (II), operably linked to a regulatory sequence; (3) a plant (IV)  
 CC comprising (III) in its genome; (4) seeds (V) obtained from (IV); and (5)  
 CC oil obtained from (V). (I) or its part can be used in antisense  
 CC inhibition or co-suppression in a transformed plant. (III) is useful for  
 CC altering the oil phenotype in a plant such as corn, soybean, wheat, rice,  
 CC canola, Brassica, sorghum, sunflower or coconut. (III) is also useful for  
 CC creating transgenic plants having altered lipid profiles. (I) can also be  
 CC used as a hybridisation probe. ACC00626 to ACC00868 and ABR40591 to  
 CC ABR40879 represent sequences used in the exemplification of the present  
 CC invention  
 XX  
 SQ Sequence 692 AA;

Query Match 100.0%; Score 42; DB 6; Length 692;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
 |||||  
 Db 493 GHHHHS 498

# RESULT 7

ABB70257

ID ABB70257 standard; protein; 1056 AA.

XX

AC ABB70257;

XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster polypeptide SEQ ID NO 37563.

XX

KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.

XX

OS Drosophila melanogaster.

XX

PN WO200171042-A2.

XX

PD 27-SEP-2001.

XX

PF 23-MAR-2001; 2001WO-US009231.

XX

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX

PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI; 2001-656860/75.  
 DR N-PSDB; ABL14360.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 XX  
 PS Disclosure; SEQ ID NO 37563; 2lpp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 1056 AA;

Query Match 100.0%; Score 42; DB 4; Length 1056;  
 Best Local Similarity 100.0%; Pred. No. 37;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
 |||||  
 Db 969 GHHHHS 974

RESULT 8  
 AAG59221  
 ID AAG59221 standard; protein; 68 AA.  
 XX  
 AC AAG59221;  
 XX  
 DT 18-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 76579.  
 XX  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-00301439.  
 XX



PR	25-FEB-1999;	99US-0121825P.
PR	05-MAR-1999;	99US-0123180P.
PR	09-MAR-1999;	99US-0123548P.
PR	23-MAR-1999;	99US-0125788P.
PR	25-MAR-1999;	99US-0126264P.
PR	29-MAR-1999;	99US-0126785P.
PR	01-APR-1999;	99US-0127462P.
PR	06-APR-1999;	99US-0128234P.
PR	08-APR-1999;	99US-0128714P.
PR	16-APR-1999;	99US-0129845P.
PR	19-APR-1999;	99US-0130077P.
PR	21-APR-1999;	99US-0130449P.
PR	23-APR-1999;	99US-0130510P.
PR	23-APR-1999;	99US-0130891P.
PR	28-APR-1999;	99US-0131449P.
PR	30-APR-1999;	99US-0132048P.
PR	30-APR-1999;	99US-0132407P.
PR	04-MAY-1999;	99US-0132484P.
PR	05-MAY-1999;	99US-0132485P.
PR	06-MAY-1999;	99US-0132486P.
PR	06-MAY-1999;	99US-0132487P.
PR	07-MAY-1999;	99US-0132863P.
PR	11-MAY-1999;	99US-0134256P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.

PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.

PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.

PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 92.9%; Score 39; DB 3; Length 68;  
 Best Local Similarity 83.3%; Pred. No. 6.7;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
 |||||:  
 Db 27 GHHHHA 32

# RESULT 9

AAG59220

ID AAG59220 standard; protein; 69 AA.

XX

AC AAG59220;

XX

DT 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 76578.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-00301439.

XX

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.

PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR	30-APR-1999;	99US-0132407P.
PR	04-MAY-1999;	99US-0132484P.
PR	05-MAY-1999;	99US-0132485P.
PR	06-MAY-1999;	99US-0132486P.
PR	06-MAY-1999;	99US-0132487P.
PR	07-MAY-1999;	99US-0132863P.
PR	11-MAY-1999;	99US-0134256P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.

PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.

PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 92.9%; Score 39; DB 3; Length 69;

Best Local Similarity 83.3%; Pred. No. 6.8;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6

||||:

## RESULT 10

AAG26198

ID AAG26198 standard; protein; 70 AA.

XX

AC AAG26198;

XX

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 30565.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-00301439.

XX

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.

PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.

PR 07-MAY-1999; 99US-0132863P.

PR 11-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.

PR 19-MAY-1999; 99US-0134941P.

PR 20-MAY-1999; 99US-0135124P.

PR 21-MAY-1999; 99US-0135353P.

PR 24-MAY-1999; 99US-0135629P.



PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.

PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.

PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 92.9%; Score 39; DB 3; Length 70;  
 Best Local Similarity 83.3%; Pred. No. 6.9;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
 |||||:  
 Db 29 GHHHHA 34

# RESULT 11

AAG59219

ID AAG59219 standard; protein; 94 AA.

XX

AC AAG59219;

XX

DT 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 76577.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.

XX  
OS Arabidopsis thaliana.  
XX  
PN EP1033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.

PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.

PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.

PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 92.9%; Score 39; DB 3; Length 94;  
 Best Local Similarity 83.3%; Pred. No. 9.5;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
 |||||:  
 Db 53 GHHHHA 58

# RESULT 12

AAG26197

ID AAG26197 standard; protein; 95 AA.

XX

AC AAG26197;

XX

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 30564.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-00301439.

XX

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR	06-APR-1999;	99US-0128234P.
PR	08-APR-1999;	99US-0128714P.
PR	16-APR-1999;	99US-0129845P.
PR	19-APR-1999;	99US-0130077P.
PR	21-APR-1999;	99US-0130449P.
PR	23-APR-1999;	99US-0130510P.
PR	23-APR-1999;	99US-0130891P.
PR	28-APR-1999;	99US-0131449P.
PR	30-APR-1999;	99US-0132048P.
PR	30-APR-1999;	99US-0132407P.
PR	04-MAY-1999;	99US-0132484P.
PR	05-MAY-1999;	99US-0132485P.
PR	06-MAY-1999;	99US-0132486P.
PR	06-MAY-1999;	99US-0132487P.
PR	07-MAY-1999;	99US-0132863P.
PR	11-MAY-1999;	99US-0134256P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.



PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.

PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161992P.

PR 28-OCT-1999; 99US-0161993P.

PR 29-OCT-1999; 99US-0162142P.

Query Match 92.9%; Score 39; DB 3; Length 95;  
Best Local Similarity 83.3%; Pred. No. 9.6;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
|||||:  
Db 54 GHHHHA 59

RESULT 13

ABB70310

ID ABB70310 standard; protein; 239 AA.

XX

AC ABB70310;

XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster polypeptide SEQ ID NO 37722.

XX

KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.

XX

OS Drosophila melanogaster.

XX

PN WO200171042-A2.

XX

PD 27-SEP-2001.

XX

PF 23-MAR-2001; 2001WO-US009231.

XX

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX

PA (PEKE ) PE CORP NY.

XX

PI Venter JC, Adams M, Li PWD, Myers EW;

XX

DR WPI; 2001-656860/75.

DR N-PSDB; ABL14413.

XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.

XX

PS Disclosure; SEQ ID NO 37722; 21pp + Sequence Listing; English.

XX

CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
CC ABB72072). The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 239 AA;

Query Match 92.9%; Score 39; DB 4; Length 239;  
Best Local Similarity 83.3%; Pred. No. 26;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
||||:  
Db 162 GHHHHT 167

RESULT 14

AAG39403

ID AAG39403 standard; protein; 324 AA.

XX

AC AAG39403;

XX

DT 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 48748.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-00301439.

XX

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 23-APR-1999; 99US-0130891P.

PR 28-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 06-MAY-1999; 99US-0132487P.

PR	07-MAY-1999;	99US-0132863P.
PR	11-MAY-1999;	99US-0134256P.
PR	14-MAY-1999;	99US-0134218P.
PR	14-MAY-1999;	99US-0134219P.
PR	14-MAY-1999;	99US-0134221P.
PR	14-MAY-1999;	99US-0134370P.
PR	18-MAY-1999;	99US-0134768P.
PR	19-MAY-1999;	99US-0134941P.
PR	20-MAY-1999;	99US-0135124P.
PR	21-MAY-1999;	99US-0135353P.
PR	24-MAY-1999;	99US-0135629P.
PR	25-MAY-1999;	99US-0136021P.
PR	27-MAY-1999;	99US-0136392P.
PR	28-MAY-1999;	99US-0136782P.
PR	01-JUN-1999;	99US-0137222P.
PR	03-JUN-1999;	99US-0137528P.
PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.

PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.

PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.  
 PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 92.9%; Score 39; DB 3; Length 324;  
 Best Local Similarity 83.3%; Pred. No. 35;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
 |||||:  
 Db 283 GHHHHA 288

RESULT 15  
 AAG39402

ID AAG39402 standard; protein; 351 AA.  
 XX  
 AC AAG39402;  
 XX  
 DT 18-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 48747.  
 XX  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PF 25-FEB-2000; 2000EP-00301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825P.  
 PR 05-MAR-1999; 99US-0123180P.  
 PR 09-MAR-1999; 99US-0123548P.  
 PR 23-MAR-1999; 99US-0125788P.  
 PR 25-MAR-1999; 99US-0126264P.  
 PR 29-MAR-1999; 99US-0126785P.  
 PR 01-APR-1999; 99US-0127462P.  
 PR 06-APR-1999; 99US-0128234P.  
 PR 08-APR-1999; 99US-0128714P.  
 PR 16-APR-1999; 99US-0129845P.  
 PR 19-APR-1999; 99US-0130077P.  
 PR 21-APR-1999; 99US-0130449P.  
 PR 23-APR-1999; 99US-0130510P.  
 PR 23-APR-1999; 99US-0130891P.  
 PR 28-APR-1999; 99US-0131449P.  
 PR 30-APR-1999; 99US-0132048P.  
 PR 30-APR-1999; 99US-0132407P.  
 PR 04-MAY-1999; 99US-0132484P.  
 PR 05-MAY-1999; 99US-0132485P.  
 PR 06-MAY-1999; 99US-0132486P.  
 PR 06-MAY-1999; 99US-0132487P.  
 PR 07-MAY-1999; 99US-0132863P.  
 PR 11-MAY-1999; 99US-0134256P.  
 PR 14-MAY-1999; 99US-0134218P.  
 PR 14-MAY-1999; 99US-0134219P.  
 PR 14-MAY-1999; 99US-0134221P.  
 PR 14-MAY-1999; 99US-0134370P.  
 PR 18-MAY-1999; 99US-0134768P.  
 PR 19-MAY-1999; 99US-0134941P.  
 PR 20-MAY-1999; 99US-0135124P.  
 PR 21-MAY-1999; 99US-0135353P.  
 PR 24-MAY-1999; 99US-0135629P.  
 PR 25-MAY-1999; 99US-0136021P.  
 PR 27-MAY-1999; 99US-0136392P.  
 PR 28-MAY-1999; 99US-0136782P.  
 PR 01-JUN-1999; 99US-0137222P.  
 PR 03-JUN-1999; 99US-0137528P.



PR	04-JUN-1999;	99US-0137502P.
PR	07-JUN-1999;	99US-0137724P.
PR	08-JUN-1999;	99US-0138094P.
PR	10-JUN-1999;	99US-0138540P.
PR	10-JUN-1999;	99US-0138847P.
PR	14-JUN-1999;	99US-0139119P.
PR	16-JUN-1999;	99US-0139452P.
PR	16-JUN-1999;	99US-0139453P.
PR	17-JUN-1999;	99US-0139492P.
PR	18-JUN-1999;	99US-0139454P.
PR	18-JUN-1999;	99US-0139455P.
PR	18-JUN-1999;	99US-0139456P.
PR	18-JUN-1999;	99US-0139457P.
PR	18-JUN-1999;	99US-0139458P.
PR	18-JUN-1999;	99US-0139459P.
PR	18-JUN-1999;	99US-0139460P.
PR	18-JUN-1999;	99US-0139461P.
PR	18-JUN-1999;	99US-0139462P.
PR	18-JUN-1999;	99US-0139463P.
PR	18-JUN-1999;	99US-0139750P.
PR	18-JUN-1999;	99US-0139763P.
PR	21-JUN-1999;	99US-0139817P.
PR	22-JUN-1999;	99US-0139899P.
PR	23-JUN-1999;	99US-0140353P.
PR	23-JUN-1999;	99US-0140354P.
PR	24-JUN-1999;	99US-0140695P.
PR	28-JUN-1999;	99US-0140823P.
PR	29-JUN-1999;	99US-0140991P.
PR	30-JUN-1999;	99US-0141287P.
PR	01-JUL-1999;	99US-0141842P.
PR	01-JUL-1999;	99US-0142154P.
PR	02-JUL-1999;	99US-0142055P.
PR	06-JUL-1999;	99US-0142390P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	19-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	22-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.

PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.

PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 14-OCT-1999; 99US-0159638P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 92.9%; Score 39; DB 3; Length 351;  
 Best Local Similarity 83.3%; Pred. No. 39;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
 |||||:  
 Db 310 GHHHHA 315

Search completed: March 5, 2004, 16:22:49  
 Job time : 8 secs

OM protein - protein search, using sw model

Run on: March 5, 2004, 16:17:14 ; Search time 1.61111 Seconds  
(without alignments)  
192.262 Million cell updates/sec

Title: US-10-057-890A-15  
Perfect score: 42  
Sequence: 1 GHHHHS 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query			ID	Description
		Match	Length	DB		
1	39	92.9	379	4	US-09-186-276B-46	Sequence 46, Appl
2	39	92.9	379	4	US-08-842-445-46	Sequence 46, Appl
3	39	92.9	379	4	US-09-186-188B-46	Sequence 46, Appl
4	39	92.9	1061	4	US-09-252-991A-23691	Sequence 23691, A
5	39	92.9	1484	2	US-08-231-193A-56	Sequence 56, Appl
6	39	92.9	1484	2	US-08-486-273A-56	Sequence 56, Appl
7	39	92.9	1484	3	US-08-940-086A-56	Sequence 56, Appl
8	39	92.9	1484	4	US-08-940-035A-56	Sequence 56, Appl
9	39	92.9	1484	4	US-08-935-105A-56	Sequence 56, Appl
10	39	92.9	1484	4	US-08-264-578-2	Sequence 2, Appli
11	39	92.9	1484	4	US-09-648-797-56	Sequence 56, Appl

12	39	92.9	1484	4	US-09-386-123-56	Sequence 56, Appl
13	38	90.5	9	1	US-08-155-171B-4	Sequence 4, Appli
14	38	90.5	9	2	US-08-435-998-4	Sequence 4, Appli
15	38	90.5	10	1	US-07-807-529A-73	Sequence 73, Appl
16	38	90.5	10	2	US-08-482-142-150	Sequence 150, App
17	38	90.5	10	2	US-08-478-572-150	Sequence 150, App
18	38	90.5	10	3	US-08-300-928C-88	Sequence 88, Appl
19	38	90.5	10	3	US-08-430-944D-88	Sequence 88, Appl
20	38	90.5	10	3	US-08-430-014-88	Sequence 88, Appl
21	38	90.5	10	3	US-08-431-184-88	Sequence 88, Appl
22	38	90.5	10	3	US-08-163-919A-17	Sequence 17, Appl
23	38	90.5	10	3	US-08-484-296-150	Sequence 150, App
24	38	90.5	10	5	PCT-US94-14073-17	Sequence 17, Appl
25	38	90.5	11	4	US-09-814-569-2	Sequence 2, Appli
26	38	90.5	13	4	US-09-418-785-3	Sequence 3, Appli
27	38	90.5	14	1	US-07-807-529A-76	Sequence 76, Appl
28	38	90.5	14	3	US-08-300-928C-91	Sequence 91, Appl
29	38	90.5	14	3	US-08-430-944D-91	Sequence 91, Appl
30	38	90.5	14	3	US-08-430-014-91	Sequence 91, Appl
31	38	90.5	14	3	US-08-431-184-91	Sequence 91, Appl
32	38	90.5	14	4	US-09-623-326-16	Sequence 16, Appl
33	38	90.5	15	2	US-08-467-603-53	Sequence 53, Appl
34	38	90.5	15	2	US-08-466-793-53	Sequence 53, Appl
35	38	90.5	15	2	US-08-491-861A-53	Sequence 53, Appl
36	38	90.5	15	4	US-09-374-671A-53	Sequence 53, Appl
37	38	90.5	17	1	US-08-155-171B-37	Sequence 37, Appl
38	38	90.5	17	2	US-08-435-998-37	Sequence 37, Appl
39	38	90.5	20	4	US-09-674-677-34	Sequence 34, Appl
40	38	90.5	21	1	US-07-927-071-3	Sequence 3, Appli
41	38	90.5	21	2	US-08-651-818A-21	Sequence 21, Appl
42	38	90.5	21	3	US-09-184-826-21	Sequence 21, Appl
43	38	90.5	22	3	US-08-256-747C-30	Sequence 30, Appl
44	38	90.5	22	3	US-08-834-130A-30	Sequence 30, Appl
45	38	90.5	22	4	US-09-660-742-3	Sequence 3, Appli

#### ALIGNMENTS

##### RESULT 1

US-09-186-276B-46

; Sequence 46, Application US/09186276B

; Patent No. 6388173

; GENERAL INFORMATION:

; APPLICANT: Benfey, Philip

; APPLICANT: DiLaurenzio, Laura

; APPLICANT: Wysocka-Diller, Joanna

; APPLICANT: Malamy, Jocelyn E.

; APPLICANT: Pysh, Leonard

; APPLICANT: Helariutta, Yrjo

; TITLE OF INVENTION: Scarecrow Gene, Promoter and Uses Thereof

; FILE REFERENCE: 5914-075-999

; CURRENT APPLICATION NUMBER: US/09/186,276B

; CURRENT FILING DATE: 1998-11-05

; PRIOR APPLICATION NUMBER: 08/842,445

; PRIOR FILING DATE: 1997-04-24

; PRIOR APPLICATION NUMBER: 08/638,617

; PRIOR FILING DATE: 1996-04-26  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 46  
; LENGTH: 379  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(379)  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-09-186-276B-46

Query Match 92.9%; Score 39; DB 4; Length 379;  
Best Local Similarity 83.3%; Pred. No. 14;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
| | | | | :  
Db 6 GHHHHT 11

#### RESULT 2

US-08-842-445-46  
; Sequence 46, Application US/08842445A  
; Patent No. 6441270  
; GENERAL INFORMATION:  
; APPLICANT: Benfey et al.  
; TITLE OF INVENTION: Scarecrow Gene, Promoter and Uses  
; TITLE OF INVENTION: Thereof  
; FILE REFERENCE: 5914-056-999  
; CURRENT APPLICATION NUMBER: US/08/842,445A  
; CURRENT FILING DATE: 1997-04-24  
; EARLIER APPLICATION NUMBER: 08/638,617  
; EARLIER FILING DATE: 1996-04-26  
; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 46  
; LENGTH: 379  
; TYPE: PRT  
; ORGANISM: Plant  
US-08-842-445-46

Query Match 92.9%; Score 39; DB 4; Length 379;  
Best Local Similarity 83.3%; Pred. No. 14;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
| | | | | :  
Db 6 GHHHHT 11

#### RESULT 3

US-09-186-188B-46  
; Sequence 46, Application US/09186188B  
; Patent No. 6455672  
; GENERAL INFORMATION:

```

; APPLICANT: Benfey et al.
; TITLE OF INVENTION: Scarecrow Gene, Promoter and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 5914-074-999
; CURRENT APPLICATION NUMBER: US/09/186,188B
; CURRENT FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 08/842,445
; PRIOR FILING DATE: 1997-04-24
; PRIOR APPLICATION NUMBER: 08/638,617
; PRIOR FILING DATE: 1996-04-26
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 46
;   LENGTH: 379
;   TYPE: PRT
;   ORGANISM: Plant
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (1)...(379)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-09-186-188B-46

```

```

Query Match          92.9%; Score 39; DB 4; Length 379;
Best Local Similarity 83.3%; Pred. No. 14;
Matches      5; Conservative 1; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      1 GHHHHS 6
        |||||:
Db      6 GHHHHT 11

```

#### RESULT 4

```

US-09-252-991A-23691
; Sequence 23691, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23691
;   LENGTH: 1061
;   TYPE: PRT
;   ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23691

```

```

Query Match          92.9%; Score 39; DB 4; Length 1061;
Best Local Similarity 83.3%; Pred. No. 39;
Matches      5; Conservative 1; Mismatches      0; Indels      0; Gaps      0;

```

Qy 1 GHHHHS 6  
 |||||:  
Db 392 GHHHHA 397

RESULT 5

US-08-231-193A-56

; Sequence 56, Application US/08231193A  
; Patent No. 5849895  
; GENERAL INFORMATION:  
; APPLICANT: Daggett, Lorrie P.  
; APPLICANT: Ellis, Steven B.  
; APPLICANT: Liaw, Chen W.  
; APPLICANT: Lu, Chin-Chun  
; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR  
; TITLE OF INVENTION: SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND USES  
THEREFOR  
; NUMBER OF SEQUENCES: 63  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Brown, Martin, Haller & McClain  
; STREET: 1660 Union Street  
; CITY: San Diego  
; STATE: CA  
; COUNTRY: U.S.A.  
; ZIP: 92101-2926  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/231,193A  
; FILING DATE: 20-APR-1994  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/052,459  
; FILING DATE: 20-APR-1993  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Seidman, Stephanie  
; REGISTRATION NUMBER: 33,779  
; REFERENCE/DOCKET NUMBER: 6362-9383  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-238-0999  
; TELEFAX: 619-238-0062  
; INFORMATION FOR SEQ ID NO: 56:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1484 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-231-193A-56

Query Match 92.9%; Score 39; DB 2; Length 1484;  
Best Local Similarity 83.3%; Pred. No. 55;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;



Qy 1 GHHHHS 6  
||||:  
Db 1360 GHHHHN 1365

RESULT 6

US-08-486-273A-56

; Sequence 56, Application US/08486273A

; Patent No. 5985586

; GENERAL INFORMATION:

; APPLICANT: Daggett, Lorrie P.

; APPLICANT: Ellis, Steven B.

; APPLICANT: Liaw, Chen W.

; APPLICANT: Lu, Chin-Chun

; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTRO SUBUNITS, DNA

; TITLE OF INVENTION: ENCODING SAME AND USES THEREFOR

; NUMBER OF SEQUENCES: 63

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Brown, Martin, Haller & McClain

; STREET: 1660 Union Street

; CITY: San Diego

; STATE: CA

; COUNTRY: U.S.A.

; ZIP: 92101-2926

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/486,273A

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/231,193

; FILING DATE: 20-APR-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Seidman, Stephanie

; REGISTRATION NUMBER: 33,779

; REFERENCE/DOCKET NUMBER: 6362-9383B

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 619-238-0999

; TELEFAX: 619-238-0062

; INFORMATION FOR SEQ ID NO: 56:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1484 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-486-273A-56

Query Match 92.9%; Score 39; DB 2; Length 1484;

Best Local Similarity 83.3%; Pred. No. 55;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
||||:  
Db 1360 GHHHHN 1365

RESULT 7

US-08-940-086A-56

; Sequence 56, Application US/08940086A

; Patent No. 6111091

; GENERAL INFORMATION:

; APPLICANT: Daggett, Lorrie P.

; APPLICANT: Ellis, Steven B.

; APPLICANT: Liaw, Chen W.

; APPLICANT: Lu, Chin-Chun

; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR

; TITLE OF INVENTION: SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND USES

THEREFOR

; NUMBER OF SEQUENCES: 63

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Heller Ehrman White & McAuliffe

; STREET: 4250 Executive Square, 7th Floor

; CITY: La Jolla

; STATE: CA

; COUNTRY: USA

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/940,086A

; FILING DATE: 29-SEPT-97

; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/231,193

; FILING DATE: 20-APR-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/052,449

; FILING DATE: 20-APR-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Seidman, Stephanie

; REGISTRATION NUMBER: 33,779

; REFERENCE/DOCKET NUMBER: 24735-9383C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 450-8400

; TELEFAX: (619) 450-8499

; INFORMATION FOR SEQ ID NO: 56:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1484 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-940-086A-56

Query Match 92.9%; Score 39; DB 3; Length 1484;  
Best Local Similarity 83.3%; Pred. No. 55;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
|||||:  
Db 1360 GHHHHN 1365

RESULT 8

US-08-940-035A-56

; Sequence 56, Application US/08940035A

; Patent No. 6316611

; GENERAL INFORMATION:

; APPLICANT: Daggett, Lorrie P.

; APPLICANT: Ellis, Steven B.

; APPLICANT: Liaw, Chen W.

; APPLICANT: Lu, Chin-Chun

; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR

; TITLE OF INVENTION: SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND USES

THEREFOR

; NUMBER OF SEQUENCES: 63

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Heller Ehrman White & McAuliffe

; STREET: 4250 Executive Square, 7th Floor

; CITY: La Jolla

; STATE: CA

; COUNTRY: U.S.A.

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/940,035A

; FILING DATE: 29-SEPT-97

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/231,193

; FILING DATE: 20-APR-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/052,449

; FILING DATE: 20-APR-1993

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Seidman, Stephanie

; REGISTRATION NUMBER: 33,779

; REFERENCE/DOCKET NUMBER: 6362-9383E

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 619-238-0999

; TELEFAX: 619-238-0062

; INFORMATION FOR SEQ ID NO: 56:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1484 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-940-035A-56

Query Match 92.9%; Score 39; DB 4; Length 1484;  
Best Local Similarity 83.3%; Pred. No. 55;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
|||||:  
Db 1360 GHHHHN 1365

RESULT 9

US-08-935-105A-56

; Sequence 56, Application US/08935105A

; Patent No. 6376660

; GENERAL INFORMATION:

; APPLICANT: Daggett, Lorrie P.

; APPLICANT: Lu, Chin-Chun

; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR

; TITLE OF INVENTION: SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND USES

THEREFOR

; NUMBER OF SEQUENCES: 63

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Heller Ehrman White & McAuliffe

; STREET: 4250 Executive Square, 7th Floor

; CITY: La Jolla

; STATE: CA

; COUNTRY: U.S.A.

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/935,105A

; FILING DATE: 29-SEPT-97

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/231,193

; FILING DATE: 20-APR-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/052,449

; FILING DATE: 20-APR-1993

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Seidman, Stephanie

; REGISTRATION NUMBER: 33,779

; REFERENCE/DOCKET NUMBER: 6362-9383D

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 619-238-0999

; TELEFAX: 619-238-0062

; INFORMATION FOR SEQ ID NO: 56:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1484 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-935-105A-56

Query Match 92.9%; Score 39; DB 4; Length 1484;  
Best Local Similarity 83.3%; Pred. No. 55;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHHHS 6  
|||||:  
Db 1360 GHHHHN 1365

RESULT 10

US-08-264-578-2

; Sequence 2, Application US/08264578

; Patent No. 6391566

; GENERAL INFORMATION:

; APPLICANT: FOLDES, Robert L.

; APPLICANT: ADAMS, Sally-Lin

; APPLICANT: KAMBOJ, Rajender

; APPLICANT: DUNCAN, H. Scott

; TITLE OF INVENTION: Modulatory Proteins of Human CNS

; TITLE OF INVENTION: Receptors

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 3000 K Street, N.W., Suite 500

; CITY: Washington, D.C.

; COUNTRY: USA

; ZIP: 20007-5109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/264,578

; FILING DATE: 23-JUN-1994

; CLASSIFICATION: 436

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/987,953

; FILING DATE: 11-DEC-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: BENT, Stephen A.

; REGISTRATION NUMBER: 29,768

; REFERENCE/DOCKET NUMBER: 16777/261/ALLE

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)672-5300

; TELEFAX: (202)672-5399

; TELEX: 904136

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1484 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-264-578-2

Query Match 92.9%; Score 39; DB 4; Length 1484;  
Best Local Similarity 83.3%; Pred. No. 55;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
|||||:  
Db 1360 GHHHHN 1365

RESULT 11

US-09-648-797-56

; Sequence 56, Application US/09648797

; Patent No. 6469142

; GENERAL INFORMATION:

; APPLICANT: Daggett, Lorrie P.

; Ellis, Steven B.

; Liaw, Chen W.

; Lu, Chin-Chun

; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR

; SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND USES

THEREFOR

; NUMBER OF SEQUENCES: 63

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Heller Ehrman White & McAuliffe

; STREET: 4250 Executive Square, 7th Floor

; CITY: La Jolla

; STATE: CA

; COUNTRY: USA

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/648,797

; FILING DATE: 28-Aug-2000

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/940,086A

; FILING DATE: 29-SEPT-97

; APPLICATION NUMBER: US 08/231,193

; FILING DATE: 20-APR-1994

; APPLICATION NUMBER: US 08/052,449

; FILING DATE: 20-APR-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Seidman, Stephanie

; REGISTRATION NUMBER: 33,779

; REFERENCE/DOCKET NUMBER: 24735-9383C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 450-8400

; TELEFAX: (619) 450-8499

; INFORMATION FOR SEQ ID NO: 56:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1484 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 56:

US-09-648-797-56

Query Match 92.9%; Score 39; DB 4; Length 1484;  
Best Local Similarity 83.3%; Pred. No. 55;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
| | | | | :  
Db 1360 GHHHHN 1365

RESULT 12

US-09-386-123-56

; Sequence 56, Application US/09386123

; Patent No. 6521413

; GENERAL INFORMATION:

; APPLICANT: Daggett, Lorrie P.

; APPLICANT: Ellis, Steven B.

; APPLICANT: Liaw, Chen W.

; APPLICANT: Lu, Chin-Chun

; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR

; TITLE OF INVENTION: SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND USES

THEREFOR

; NUMBER OF SEQUENCES: 63

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Heller Ehrman White & McAuliffe

; STREET: 4250 Executive Square, 7th Floor

; CITY: La Jolla

; STATE: CA

; COUNTRY: U.S.A.

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/386,123

; FILING DATE:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/486,273

; FILING DATE: 06-JUNE-95

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/231,193

; FILING DATE: 20-APR-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/052,449

; FILING DATE: 20-APR-1993

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Seidman, Stephanie

; REGISTRATION NUMBER: 33,779

; REFERENCE/DOCKET NUMBER: 6362-9383F

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 858-450-8403

; TELEFAX: 858-587-5360

; INFORMATION FOR SEQ ID NO: 56:

; SEQUENCE CHARACTERISTICS:

;     LENGTH: 1484 amino acids  
;     TYPE: amino acid  
;     TOPOLOGY: linear  
;     MOLECULE TYPE: protein  
US-09-386-123-56

Query Match                   92.9%; Score 39; DB 4; Length 1484;  
Best Local Similarity   83.3%; Pred. No. 55;  
Matches       5; Conservative   1; Mismatches       0; Indels       0; Gaps       0;

Qy            1 GHHHHS 6  
              |||||:  
Db           1360 GHHHHN 1365

RESULT 13

US-08-155-171B-4

; Sequence 4, Application US/08155171B  
; Patent No. 5543264

; GENERAL INFORMATION:

; APPLICANT: Anderson, Carl W.  
; APPLICANT: Mangel, Walter F.  
; TITLE OF INVENTION: Co-Factor Activated Recombinant  
; TITLE OF INVENTION: Adenovirus Proteinases (As Amended)  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02173

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/155,171B  
; FILING DATE: 19-NOV-1993  
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/851,217  
; FILING DATE: 13-MAR-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/545,585  
; FILING DATE: 29-JUN-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: Granahan, Patricia  
; REGISTRATION NUMBER: 32,227  
; REFERENCE/DOCKET NUMBER: BNL91-01A2, AUI93-22

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 861-6240  
; TELEFAX: (617) 861-9540

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids



; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
US-08-155-171B-4

Query Match 90.5%; Score 38; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 3e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5  
| | | | |  
Db 1 GHHHH 5

RESULT 14

US-08-435-998-4

; Sequence 4, Application US/08435998  
; Patent No. 5935840  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, Carl W.  
; APPLICANT: Mangel, Walter F.  
; TITLE OF INVENTION: Co-Factor Activated Recombinant  
; TITLE OF INVENTION: Adenovirus Proteinases (As Amended)  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02173  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/435,998  
; FILING DATE: 05-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/155,171  
; FILING DATE: 19-NOV-1993  
; APPLICATION NUMBER: US 07/851,217  
; FILING DATE: 13-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/545,585  
; FILING DATE: 29-JUN-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Granahan, Patricia  
; REGISTRATION NUMBER: 32,227  
; REFERENCE/DOCKET NUMBER: BNL91-01A2, AUI93-22  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 861-6240  
; TELEFAX: (617) 861-9540  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:

;     LENGTH: 9 amino acids  
;     TYPE: amino acid  
;     STRANDEDNESS:  
;     TOPOLOGY: linear  
US-08-435-998-4

Query Match                   90.5%; Score 38; DB 2; Length 9;  
Best Local Similarity   100.0%; Pred. No. 3e+05;  
Matches       5; Conservative   0; Mismatches    0; Indels       0; Gaps       0;

Qy            1 GHHHH 5  
              |||||  
Db            1 GHHHH 5

RESULT 15

US-07-807-529A-73

; Sequence 73, Application US/07807529A  
; Patent No. 5547669

; GENERAL INFORMATION:

; APPLICANT: Rogers, Bruce L.  
; APPLICANT: Morgenstern, Jay  
; APPLICANT: Bond, Julian F.  
; APPLICANT: Garman, Richard D.  
; APPLICANT: Greenstein, Julia L.  
; APPLICANT: Kuo, Mei-chang  
; APPLICANT: Morvile, Malcolm

; TITLE OF INVENTION: RECOMBITOPE PEPTIDES  
; NUMBER OF SEQUENCES: 76

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION  
; STREET: One Kendall Square, Building 600  
; CITY: Cambridge  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02139

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII TEXT

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/807,529A  
; FILING DATE: 19911213  
; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/662,276  
; FILING DATE: 28-FEB-1991  
; APPLICATION NUMBER: US 07/431,565  
; FILING DATE: 03-NOV-1989

; ATTORNEY/AGENT INFORMATION:

; NAME: Channing, Stacey L.  
; REGISTRATION NUMBER: 31,095  
; REFERENCE/DOCKET NUMBER: IPC-027/imi-015

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 494-0060  
; INFORMATION FOR SEQ ID NO: 73:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: internal  
US-07-807-529A-73

Query Match 90.5%; Score 38; DB 1; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.48;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5  
| | | | |  
Db 2 GHHHH 6

Search completed: March 5, 2004, 16:30:37  
Job time : 2.61111 secs

OM protein - protein search, using sw model

Run on: March 5, 2004, 16:16:19 ; Search time 1.37037 Seconds  
 (without alignments)  
 421.163 Million cell updates/sec

Title: US-10-057-890A-15  
 Perfect score: 42  
 Sequence: 1 GHHHHS 6  
 Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : PIR\_78:\*  
 1: pirl:\*  
 2: pir2:\*  
 3: pir3:\*  
 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	39	92.9	140	2	T27059	hypothetical prote
2	39	92.9	361	2	F87286	cation efflux fami
3	39	92.9	378	2	T51237	scarecrow-like pro
4	39	92.9	558	2	T01343	hypothetical prote
5	39	92.9	1484	2	S52086	N-methyl-D-asparta
6	38	90.5	60	2	C64698	probable histidine
7	38	90.5	77	2	T16436	hypothetical prote
8	38	90.5	83	2	T16435	hypothetical prote
9	38	90.5	102	2	T30119	hypothetical prote
10	38	90.5	107	2	A29995	protamine P2 precu
11	38	90.5	114	2	S37150	asr2 protein - tom
12	38	90.5	115	2	A23925	proline-rich phosph
13	38	90.5	118	1	A31429	hisactophilin [val

14	38	90.5	128	1	H69052	conserved hypothet
15	38	90.5	130	2	S14983	extensin class I (
16	38	90.5	131	2	A53024	Leydig insulin-lik
17	38	90.5	132	2	G75466	probable ferric up
18	38	90.5	148	2	G83681	urease accessory p
19	38	90.5	160	2	A81075	FKBP-type peptidyl
20	38	90.5	161	2	E43719	ureE protein - Pro
21	38	90.5	161	2	H85120	probable proline-r
22	38	90.5	191	2	AD2871	urease accessory p
23	38	90.5	192	2	T39367	hypothetical prote
24	38	90.5	199	2	T48099	hypothetical prote
25	38	90.5	219	2	T10563	proline-rich prote
26	38	90.5	222	2	F97647	urease accessory p
27	38	90.5	235	2	D89101	protein F25E5.8 [i
28	38	90.5	274	2	A64978	hypothetical prote
29	38	90.5	290	2	T49631	probable Ni-bindin
30	38	90.5	294	2	E69759	hypothetical prote
31	38	90.5	298	2	S41469	homeotic protein M
32	38	90.5	299	2	S32874	hypB protein - Rhi
33	38	90.5	302	2	A55641	homeotic protein G
34	38	90.5	303	2	A56837	homeotic protein M
35	38	90.5	303	2	B49122	homeobox protein M
36	38	90.5	303	2	A48130	growth arrest-spec
37	38	90.5	305	2	JN0647	hydrogenase expres
38	38	90.5	306	2	T44684	hypothetical prote
39	38	90.5	309	2	T29293	hypothetical prote
40	38	90.5	323	2	I51751	homeotic protein o
41	38	90.5	336	1	S75947	hypothetical prote
42	38	90.5	343	2	T29547	hypothetical prote
43	38	90.5	348	2	T06385	probable Fe(II) tr
44	38	90.5	349	2	S61414	DNA-binding protei
45	38	90.5	351	1	KGZQHL	histidine-rich gly

# ALIGNMENTS

## RESULT 1

T27059

hypothetical protein Y51A2A.6 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C;Accession: T27059

R;McMurray, A.

submitted to the EMBL Data Library, October 1998

A;Reference number: Z20304

A;Accession: T27059

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-140 <WIL>

A;Cross-references: EMBL:AL032635; PIDN:CAA21601.1; GSPDB:GN00023; CESP:Y51A2A.6

A;Experimental source: clone Y51A2A

C;Genetics:

A;Gene: CESP:Y51A2A.6

A;Map position: 5

A;Introns: 93/3; 129/1

Query Match 92.9%; Score 39; DB 2; Length 140;  
Best Local Similarity 83.3%; Pred. No. 8.5;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
|||||:  
Db 53 GHHHHN 58

RESULT 2

F87286

cation efflux family protein [imported] - *Caulobacter crescentus*

C;Species: *Caulobacter crescentus*

C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001

C;Accession: F87286

R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.F.; Alley, M.; Ohta, N.; Maddock, J.R.; Potocka, I.; Nelson, W.C.; Newton, A.; Stephens, C.; Phadke, N.D.; Ely, B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonay, J.F.; Smit, J.; Craven, M.; Khouri, H.; Shetty, J.; Berry, K.; Utterback, T.; Tran, K.; Wolf, A.; Vamathevan, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of *Caulobacter crescentus*.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: F87286

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-361 <STO>

A;Cross-references: GB:AE005673; NID:gl3421446; PIDN:AAK22290.1; GSPDB:GN00148

C;Genetics:

A;Gene: CC0303

Query Match 92.9%; Score 39; DB 2; Length 361;  
Best Local Similarity 83.3%; Pred. No. 22;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
|||||:  
Db 64 GHHHHA 69

RESULT 3

T51237

scarecrow-like protein 6 [imported] - *Arabidopsis thaliana* (fragment)

C;Species: *Arabidopsis thaliana* (mouse-ear cress)

C;Date: 28-Jul-2000 #sequence\_revision 28-Jul-2000 #text\_change 28-Jul-2000

C;Accession: T51237

R;Pysh, L.D.; Wysocka-Diller, J.W.; Camilleri, C.; Bouchez, D.; Benfey, P.N. Plant J. 18, 111-119, 1999

A;Title: The GRAS gene family in *Arabidopsis*: sequence characterization and basic expression analysis of the SCARECROW-LIKE genes.

A;Reference number: Z25337; MUID:99272994; PMID:10341448

A;Accession: T51237

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: mRNA

A;Residues: 1-378 <PYS>

A;Cross-references: EMBL:AF036303; PIDN:AAD24406.1  
C;Genetics:  
A;Gene: SCL6  
A;Map position: 4

Query Match 92.9%; Score 39; DB 2; Length 378;  
Best Local Similarity 83.3%; Pred. No. 23;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
|||||:  
Db 6 GHHHHT 11

RESULT 4

T01343

hypothetical protein F6N15.20 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 14-May-1999

C;Accession: T01343

R;Ryan, E.; Edwards, J.; Pape, K.

submitted to the EMBL Data Library, May 1998

A;Description: The sequence of A. thaliana F6N15.

A;Reference number: Z14297

A;Accession: T01343

A;Status: translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-558 <RYA>

A;Cross-references: EMBL:AF069299; NID:g3193311; PID:g3193314

A;Experimental source: cultivar Columbia

C;Genetics:

A;Map position: 4

A;Note: F6N15.20

Query Match 92.9%; Score 39; DB 2; Length 558;  
Best Local Similarity 83.3%; Pred. No. 35;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
|||||:  
Db 186 GHHHHT 191

RESULT 5

S52086

N-methyl-D-aspartate receptor chain NR3 - human

C;Species: Homo sapiens (man)

C;Date: 28-Oct-1996 #sequence\_revision 07-Feb-1997 #text\_change 21-Jan-2000

C;Accession: S52086; S70925

R;Adams, S.L.; Foldes, R.L.; Kamboj, R.K.

Biochim. Biophys. Acta 1260, 105-108, 1995

A;Title: Human N-methyl-D-aspartate receptor modulatory subunit hNR3: cloning and sequencing of the cDNA and primary structure of the protein.

A;Reference number: S52086; MUID:95092783; PMID:7999784

A;Accession: S52086

A;Molecule type: mRNA

A;Residues: 1-1484 <ADA>

A;Cross-references: EMBL:U11287  
 A;Note: 407-Asn was also found  
 R;Foldes, R.L.  
 submitted to the EMBL Data Library, June 1994  
 A;Reference number: S70925  
 A;Accession: S70925  
 A;Molecule type: mRNA  
 A;Residues: 1-270,'A',272-919,'RP',922-1484 <FOL>  
 A;Cross-references: EMBL:U11287; NID:g560546; PIDN:AAB60368.1; PID:g560547  
 C;Superfamily: N-methyl-D-aspartate receptor 2A; glutamate receptor homology  
 C;Keywords: ion channel; neurotransmitter receptor; transmembrane protein  
 F;428-855/Domain: glutamate receptor homology <GRH>

Query Match 92.9%; Score 39; DB 2; Length 1484;  
 Best Local Similarity 83.3%; Pred. No. 93;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
 |||||:  
 Db 1360 GHHHHN 1365

# RESULT 6

C64698  
 probable histidine-rich metal-binding protein - *Helicobacter pylori*  
 C;Species: *Helicobacter pylori*  
 A;Variety: strains J99, 26695  
 C;Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999  
 C;Accession: C64698; C71821  
 R;Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.;  
 Fleischmann, R.D.; Ketchum, K.A.; Klenk, H.P.; Gill, S.; Dougherty, B.A.;  
 Nelson, K.; Quackenbush, J.; Zhou, L.; Kirkness, E.F.; Peterson, S.; Loftus, B.;  
 Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, K.; Fitzegerald,  
 L.M.; Lee, N.; Adams, M.D.; Hickey, E.K.; Berg, D.E.; Gocayne, J.D.; Utterback,  
 T.R.; Peterson, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.;  
 Bowman, C.; Watthey, L.  
 Nature 388, 539-547, 1997  
 A;Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.;  
 Fraser, C.M.; Venter, J.C.  
 A;Title: The complete genome sequence of the gastric pathogen *Helicobacter*  
*pylori*.  
 A;Reference number: A64520; MUID:97394467; PMID:9252185  
 A;Accession: C64698  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-60 <TOM>  
 A;Cross-references: GB:AE000643; GB:AE000511; NID:g2314598; PIDN:AAD08471.1;  
 PID:g2314604; TIGR:HP1427  
 A;Experimental source: strain 26695  
 R;Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.;  
 Smith, D.R.; Noonan, B.; Guild, B.C.; deJonge, B.L.; Carmel, G.; Tummino, P.J.;  
 Caruso, A.; Uria-Nickelsen, M.; Mills, D.M.; Ives, C.; Gibson, R.; Merberg, D.;  
 Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Trust, T.J.  
 Nature 397, 176-180, 1999  
 A;Title: Genomic sequence comparison of two unrelated isolates of the human  
 gastric pathogen *Helicobacter pylori*.  
 A;Reference number: A71800; MUID:99120557; PMID:9923682



A;Accession: C71821  
A;Molecule type: DNA  
A;Residues: 1-60 <ARN>  
A;Cross-references: GB:AE001555; GB:AE001439; NID:g4155929; PIDN:AAD06898.1;  
PID:g4155931  
A;Experimental source: strain J99  
C;Genetics:  
A;Gene: HP1427; jhp1320

Query Match 90.5%; Score 38; DB 2; Length 60;  
Best Local Similarity 100.0%; Pred. No. 5.2;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5  
|||||  
Db 10 GHHHH 14

RESULT 7

T16436  
hypothetical protein F53A9.1 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C;Accession: T16436  
R;Miller, N.  
submitted to the EMBL Data Library, March 1995  
A;Description: The sequence of C. elegans cosmid F53A9.  
A;Reference number: Z18513  
A;Accession: T16436  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-77 <MIL>  
A;Cross-references: EMBL:U23523; NID:g746551; PID:g746552; PIDN:AAC46556.1;  
CESP:F53A9.1  
A;Experimental source: strain Bristol N2  
C;Genetics:  
A;Gene: CESP:F53A9.1

Query Match 90.5%; Score 38; DB 2; Length 77;  
Best Local Similarity 100.0%; Pred. No. 6.7;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5  
|||||  
Db 45 GHHHH 49

RESULT 8

T16435  
hypothetical protein F53A9.2 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C;Accession: T16435  
R;Miller, N.  
submitted to the EMBL Data Library, March 1995  
A;Description: The sequence of C. elegans cosmid F53A9.  
A;Reference number: Z18513

A;Accession: T16435  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-83 <MIL>  
A;Cross-references: EMBL:U23523; NID:g746551; PID:g746553; PIDN:AAC46557.1;  
CESP:F53A9.2  
A;Experimental source: strain Bristol N2  
C;Genetics:  
A;Gene: CESP:F53A9.2

Query Match 90.5%; Score 38; DB 2; Length 83;  
Best Local Similarity 100.0%; Pred. No. 7.2;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5  
|||||  
Db 44 GHHHH 48

RESULT 9

T30119

hypothetical protein F22H10.2 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000

C;Accession: T30119

R;Langston, Y.; Hawkins, J.

submitted to the EMBL Data Library, September 1996

A;Description: The sequence of C. elegans cosmid F22H10.

A;Reference number: Z20740

A;Accession: T30119

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-102 <LAN>

A;Cross-references: EMBL:U70845; PIDN:AAB09100.1; GSPDB:GN00028; CESP:F22H10.2

A;Experimental source: strain Bristol N2; clone F22H10

C;Genetics:

A;Gene: CESP:F22H10.2

A;Map position: X

A;Introns: 16/1

Query Match 90.5%; Score 38; DB 2; Length 102;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5  
|||||  
Db 60 GHHHH 64

RESULT 10

A29995

protamine P2 precursor - mouse

N;Alternate names: sperm histone P2

C;Species: Mus musculus (house mouse)

C;Date: 09-Sep-1988 #sequence\_revision 07-Jun-1996 #text\_change 23-Jul-1999

C;Accession: A27809; A29995; S21179; S03821; S16210; I64886

R;Yelick, P.C.; Balhorn, R.; Johnson, P.A.; Corzett, M.; Mazrimas, J.A.; Kleene, K.C.; Hecht, N.B.  
Mol. Cell. Biol. 7, 2173-2179, 1987  
A;Title: Mouse protamine 2 is synthesized as a precursor whereas mouse protamine 1 is not.  
A;Reference number: A27809; MUID:87257931; PMID:3600661  
A;Accession: A27809  
A;Molecule type: mRNA; protein  
A;Residues: 1-107 <YEL>  
A;Cross-references: GB:M16456; NID:g200490; PIDN:AAA39981.1; PID:g200491  
R;Bellve, A.R.; McKay, D.J.; Renaux, B.S.; Dixon, G.H.  
Biochemistry 27, 2890-2897, 1988  
A;Title: Purification and characterization of mouse protamines P1 and P2. Amino acid sequence of P2.  
A;Reference number: A29995; MUID:88294032; PMID:3401454  
A;Accession: A29995  
A;Molecule type: protein  
A;Residues: 1-107 <BEL>  
A;Note: the signal sequence was partially sequenced  
R;Chauviere, M.; Martinage, A.; Debarle, M.; Sautiere, P.; Chevaillier, P.  
Eur. J. Biochem. 204, 759-765, 1992  
A;Title: Molecular characterization of six intermediate proteins in the processing of mouse protamine P2 precursor.  
A;Reference number: S21179; MUID:92174934; PMID:1541289  
A;Accession: S21179  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 2-42 <CHA>  
R;Johnson, P.A.; Peschon, J.J.; Yelick, P.C.; Palmiter, R.D.; Hecht, N.B.  
Biochim. Biophys. Acta 950, 45-53, 1988  
A;Title: Sequence homologies in the mouse protamine 1 and 2 genes.  
A;Reference number: S03820; MUID:88193085; PMID:3358932  
A;Accession: S03821  
A;Molecule type: DNA  
A;Residues: 1-107 <JOH>  
A;Cross-references: EMBL:X07626; NID:g53792; PIDN:CAA30473.1; PID:g53793  
R;Carre-Eusebe, D.; Lederer, F.; Le, K.H.D.; Elsevier, S.M.  
Biochem. J. 277, 39-45, 1991  
A;Title: Processing of the precursor of protamine P2 in mouse. Peptide mapping and N-terminal sequence analysis of intermediates.  
A;Reference number: S16210; MUID:91307542; PMID:1854346  
A;Accession: S16210  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 'X',22-23,'XX',26-30,'X',32-33,45-58,'X',60-66 <BIO>  
R;Hecht, N.B.  
Ann. N. Y. Acad. Sci. 513, 91-101, 1987  
A;Title: gene expression during spermatogenesis.  
A;Reference number: I51954  
A;Accession: I64886  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-107 <RES>  
A;Cross-references: GB:M27501; NID:g200504; PIDN:AAA39986.1; PID:g200505  
C;Genetics:  
A;Map position: 16  
A;Introns: 87/1

C;Superfamily: sperm histone  
C;Keywords: DNA binding; nucleus  
F;1-44/Domain: signal sequence #status experimental <SIG>  
F;45-107/Product: sperm histone P2 #status experimental <MAT>

Query Match 90.5%; Score 38; DB 2; Length 107;  
Best Local Similarity 100.0%; Pred. No. 9.3;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5  
|||||  
Db 46 GHHHH 50

RESULT 11

S37150  
asr2 protein - tomato  
C;Species: Lycopersicon esculentum (tomato)  
C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Sep-1997  
C;Accession: S37150  
R;Amitai, H.; Scolnik, P.A.; Bar-Zvi, D.  
submitted to the EMBL Data Library, September 1993  
A;Reference number: S37150  
A;Accession: S37150  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-114 <AMI>  
A;Cross-references: EMBL:X74907; NID:g400468; PID:g400469  
C;Genetics:  
A;Introns: 53/3

Query Match 90.5%; Score 38; DB 2; Length 114;  
Best Local Similarity 100.0%; Pred. No. 9.9;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5  
|||||  
Db 107 GHHHH 111

RESULT 12

A23925  
proline-rich phosphoprotein - crab-eating macaque  
C;Species: Macaca fascicularis (crab-eating macaque)  
C;Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 25-Oct-1996  
C;Accession: A23925  
R;Oppenheim, F.G.; Offner, G.D.; Troxler, R.F.  
J. Biol. Chem. 260, 10671-10679, 1985  
A;Title: Amino acid sequence of a proline-rich phosphoglycoprotein from parotid  
secretion of the subhuman primate Macaca fascicularis.  
A;Reference number: A23925; MUID:85289254; PMID:4030765  
A;Accession: A23925  
A;Molecule type: protein  
A;Residues: 1-115 <OPP>  
C;Superfamily: proline-rich protein  
C;Keywords: phosphoprotein

Query Match 90.5%; Score 38; DB 2; Length 115;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5  
|||||  
Db 80 GHHHH 84

RESULT 13

A31429

hisactophilin [validated] - slime mold (*Dictyostelium discoideum*)

N;Alternate names: histidine-rich actin-binding protein

C;Species: *Dictyostelium discoideum*

C;Date: 31-Jul-1989 #sequence\_revision 02-Aug-1994 #text\_change 15-Sep-2000

C;Accession: A31429; A30787

R;Scheel, J.; Ziegelbauer, K.; Kupke, T.; Humbel, B.M.; Noegel, A.A.; Gerisch, G.; Schleicher, M.

J. Biol. Chem. 264, 2832-2839, 1989

A;Title: Hisactophilin, a histidine-rich actin-binding protein from *Dictyostelium discoideum*.

A;Reference number: A31429; MUID:89123382; PMID:2914932

A;Accession: A31429

A;Molecule type: mRNA

A;Residues: 1-118 <SCH>

A;Cross-references: GB:J04472; NID:g167812; PIDN:AAA33218.1; PID:g167813

R;Habazettl, J.; Gondol, D.; Wiltsccheck, R.; Otlewski, J.; Schleicher, M.; Holak, T.A.

submitted to the Brookhaven Protein Data Bank, May 1994

A;Reference number: A52585; PDB:1HCD

A;Contents: annotation; conformation by (1)H-, (13)C-, and (15)N-NMR, residues 1-118

A;Note: recombinant form expressed in *Escherichia coli* includes Met-1 and lacks post-translational modifications of the mature protein

R;Habazettl, J.; Gondol, D.; Wiltsccheck, R.; Otlewski, J.; Schleicher, M.; Holak, T.A.

Nature 359, 855-858, 1992

A;Title: Structure of hisactophilin is similar to interleukin-1beta and fibroblast growth factor.

A;Reference number: A59170; MUID:93063300; PMID:1436061

A;Contents: annotation; conformation by (1)H-, (13)C-, and (15)N-NMR

R;Hanakam, F.; Gerisch, G.; Lotz, S.; Alt, T.; Seelig, A.

Biochemistry 35, 11036-11044, 1996

A;Title: Binding of hisactophilin I and II to lipid membranes is controlled by a pH-dependent myristoyl-histidine switch.

A;Reference number: A59169; MUID:96374214; PMID:8780505

A;Contents: annotation

R;Hanakam, F.; Eckerskorn, C.; Lottspeich, F.; Mueller-Taubenberger, A.; Schaefer, W.; Gerisch, G.

J. Biol. Chem. 270, 596-602, 1995

A;Title: The pH-sensitive actin-binding protein hisactophilin of *Dictyostelium* exists in two isoforms which both are myristoylated and distributed between plasma membrane and cytoplasm.

A;Reference number: A59171; MUID:95122497; PMID:7822284

A;Contents: annotation

R;Urban, M.; Gerisch, G.

unpublished results, cited by Schleicher, M., in Guidebook to the Cytoskeletal and Motor Proteins, Kreis. T. and Vale, R., eds., pp.54-55, Oxford University Press, Oxford, 1993

A;Reference number: A38915

A;Contents: annotation; palmitate binding

A;Note: one or more of the serines is phosphorylated

C;Comment: Hisactophilin binds to F-actin in a pH-dependent manner, inducing actin polymerization. It is suggested to act as an intracellular pH sensor that links chemotactic signals to responses in the microfilament system.

C;Superfamily: hisactophilin

C;Keywords: actin binding; blocked amino end; duplication; lipoprotein; myristylation; tandem repeat; thiolester bond

F;2-118/Product: hisactophilin #status experimental <MAT>

F;34-86/Region: 13-residue repeats (F-H-V-E-N-H-G-G-K-V-A-L-K)

F;2/Modified site: myristylated amino end (Gly) (in mature form) #status experimental

F;3/Modified site: aspartic acid (Asn) #status predicted

F;49/Binding site: palmitate (Cys) (covalent) (partial) #status experimental

Query Match 90.5%; Score 38; DB 1; Length 118;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5  
|||||  
Db 87 GHHHH 91

#### RESULT 14

H69052

conserved hypothetical protein MTH1397 - Methanobacterium thermoautotrophicum (strain Delta H)

C;Species: Methanobacterium thermoautotrophicum

C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000

C;Accession: H69052

R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Bashirzadeh, R.; Blakely, D.; Cook, R.; Gilbert, K.; Harrison, D.; Hoang, L.; Keagle, P.; Lumm, W.; Pothier, B.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.; Caruso, A.; Bush, D.; Safer, H.; Patwell, D.; Prabhakar, S.; McDougall, S.; Shimer, G.; Goyal, A.; Pietrokowski, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functional analysis and comparative genomics.

A;Reference number: A69000; MUID:98037514; PMID:9371463

A;Accession: H69052

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-128 <MTH>

A;Cross-references: GB:AE000902; GB:AE000666; NID:g2622500; PIDN:AAB85874.1; PID:g2622508

A;Experimental source: strain Delta H

C;Genetics:

A;Gene: MTH1397

A;Start codon: GTG

C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0970

Query Match 90.5%; Score 38; DB 1; Length 128;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5  
|||||  
Db 85 GHHHH 89

RESULT 15

S14983

extensin class I (clone w10-1 L) - tomato (fragment)

C;Species: Lycopersicon esculentum (tomato)

C;Date: 07-May-1998 #sequence\_revision 15-May-1998 #text\_change 17-Jul-1998

C;Accession: S14983

R;Showalter, A.M.; Zhou, J.; Rumeau, D.; Worst, S.G.; Varner, J.E.

Plant Mol. Biol. 16, 547-565, 1991

A;Title: Tomato extensin and extensin-like cDNAs: structure and expression in response to wounding.

A;Reference number: S14970; MUID:91329690; PMID:1714316

A;Accession: S14983

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-130 <SHO>

A;Cross-references: EMBL:X55694

A;Experimental source: cv. UC82B

C;Keywords: cell wall; glycoprotein; hydroxyproline

Query Match 90.5%; Score 38; DB 2; Length 130;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5  
|||||  
Db 9 GHHHH 13

Search completed: March 5, 2004, 16:28:56

Job time : 2.37037 secs

OM protein - protein search, using sw model

Run on: March 5, 2004, 16:22:54 ; Search time 3.24074 Seconds  
 (without alignments)  
 390.935 Million cell updates/sec

Title: US-10-057-890A-15  
 Perfect score: 42  
 Sequence: 1 GHHHHS 6  
 Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	%	Query					
No.	Score	Match	Length	DB	ID		Description



1	42	100.0	6	14	US-10-057-890A-15	Sequence 15, Appl
2	42	100.0	138	14	US-10-057-890A-10	Sequence 10, Appl
3	42	100.0	157	14	US-10-057-890A-31	Sequence 31, Appl
4	39	92.9	324	15	US-10-374-780A-1363	Sequence 1363, Ap
5	39	92.9	379	14	US-10-253-007-46	Sequence 46, Appl
6	39	92.9	1480	10	US-09-922-011-10	Sequence 10, Appl
7	39	92.9	1484	9	US-09-945-901-56	Sequence 56, Appl
8	39	92.9	1484	13	US-10-007-747-56	Sequence 56, Appl
9	39	92.9	1484	14	US-10-038-937-56	Sequence 56, Appl
10	39	92.9	1484	14	US-10-146-806-2	Sequence 2, Appli
11	38	90.5	6	14	US-10-179-784-39	Sequence 39, Appl
12	38	90.5	9	9	US-09-821-984-44	Sequence 44, Appl
13	38	90.5	9	9	US-09-284-663A-25	Sequence 25, Appl
14	38	90.5	9	9	US-09-854-280-18	Sequence 18, Appl
15	38	90.5	9	9	US-09-854-208-18	Sequence 18, Appl
16	38	90.5	9	14	US-10-203-013-27	Sequence 27, Appl
17	38	90.5	9	14	US-10-203-013-29	Sequence 29, Appl
18	38	90.5	10	8	US-08-464-363-73	Sequence 73, Appl
19	38	90.5	10	9	US-09-981-636-2	Sequence 2, Appli
20	38	90.5	10	9	US-09-981-636-3	Sequence 3, Appli
21	38	90.5	10	10	US-09-976-935-31	Sequence 31, Appl
22	38	90.5	10	11	US-09-933-780C-24	Sequence 24, Appl
23	38	90.5	10	14	US-10-104-919-60	Sequence 60, Appl
24	38	90.5	10	15	US-10-351-157-72	Sequence 72, Appl
25	38	90.5	10	16	US-10-395-741B-61	Sequence 61, Appl
26	38	90.5	11	9	US-09-814-569-2	Sequence 2, Appli
27	38	90.5	13	15	US-10-297-229-67	Sequence 67, Appl
28	38	90.5	14	8	US-08-464-363-76	Sequence 76, Appl
29	38	90.5	14	9	US-09-900-530A-19	Sequence 19, Appl
30	38	90.5	15	9	US-09-374-671-53	Sequence 53, Appl
31	38	90.5	15	14	US-10-196-107A-53	Sequence 53, Appl
32	38	90.5	19	14	US-10-342-103-17	Sequence 17, Appl
33	38	90.5	20	14	US-10-408-930-32	Sequence 32, Appl
34	38	90.5	21	10	US-09-832-464-21	Sequence 21, Appl
35	38	90.5	22	9	US-09-331-631A-16	Sequence 16, Appl
36	38	90.5	22	14	US-10-147-095-16	Sequence 16, Appl
37	38	90.5	23	15	US-10-354-774-24	Sequence 24, Appl
38	38	90.5	23	15	US-10-271-012-24	Sequence 24, Appl
39	38	90.5	24	9	US-09-934-465-8	Sequence 8, Appli
40	38	90.5	24	9	US-09-884-733-9	Sequence 9, Appli
41	38	90.5	24	13	US-10-080-455-5	Sequence 5, Appli
42	38	90.5	24	14	US-10-112-193-9	Sequence 9, Appli
43	38	90.5	26	9	US-09-773-385-18	Sequence 18, Appl
44	38	90.5	26	14	US-10-209-041-12	Sequence 12, Appl
45	38	90.5	27	14	US-10-342-103-4	Sequence 4, Appli

#### ALIGNMENTS

RESULT 1  
 US-10-057-890A-15  
 ; Sequence 15, Application US/10057890A  
 ; Publication No. US20030044901A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Coleman, Timothy

```

; APPLICANT: Mansfield, Brian
; TITLE OF INVENTION: Scaffold Fusion Polypeptides, Composition for Making the
Same, and Methods
; TITLE OF INVENTION: of Using the Same.
; FILE REFERENCE: PF537
; CURRENT APPLICATION NUMBER: US/10/057,890A
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: 60/265,782
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,858
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 15
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-890A-15

```

```

Query Match          100.0%;  Score 42;  DB 14;  Length 6;
Best Local Similarity 100.0%;  Pred. No. 7.1e+05;
Matches      6;  Conservative    0;  Mismatches    0;  Indels      0;  Gaps      0;

```

```

Qy      1 GHHHHS 6
        |||||
Db      1 GHHHHS 6

```

#### RESULT 2

```

US-10-057-890A-10
; Sequence 10, Application US/10057890A
; Publication No. US20030044901A1
; GENERAL INFORMATION:
; APPLICANT: Coleman, Timothy
; APPLICANT: Mansfield, Brian
; TITLE OF INVENTION: Scaffold Fusion Polypeptides, Composition for Making the
Same, and Methods
; TITLE OF INVENTION: of Using the Same.
; FILE REFERENCE: PF537
; CURRENT APPLICATION NUMBER: US/10/057,890A
; CURRENT FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: 60/265,782
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,858
; PRIOR FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 10
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-890A-10

```

```

Query Match          100.0%;  Score 42;  DB 14;  Length 138;
Best Local Similarity 100.0%;  Pred. No. 21;
Matches      6;  Conservative    0;  Mismatches    0;  Indels      0;  Gaps      0;

```

```

Qy      1 GHHHHS 6
        |||||

```

Db 55 GHHHHS 60

RESULT 3

US-10-057-890A-31

; Sequence 31, Application US/10057890A  
; Publication No. US20030044901A1  
; GENERAL INFORMATION:  
; APPLICANT: Coleman, Timothy  
; APPLICANT: Mansfield, Brian  
; TITLE OF INVENTION: Scaffold Fusion Polypeptides, Composition for Making the Same, and Methods  
; TITLE OF INVENTION: of Using the Same.  
; FILE REFERENCE: PF537  
; CURRENT APPLICATION NUMBER: US/10/057,890A  
; CURRENT FILING DATE: 2002-01-29  
; PRIOR APPLICATION NUMBER: 60/265,782  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/265,858  
; PRIOR FILING DATE: 2001-01-31  
; NUMBER OF SEQ ID NOS: 32  
; SEQ ID NO 31  
; LENGTH: 157  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-057-890A-31

Query Match 100.0%; Score 42; DB 14; Length 157;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
|||||  
Db 74 GHHHHS 79

RESULT 4

US-10-374-780A-1363

; Sequence 1363, Application US/10374780A  
; Publication No. US20040019927A1  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, Bradley K  
; APPLICANT: Riechmann, Jose Luis  
; APPLICANT: Jiang, Cai-Zhong  
; APPLICANT: Heard, Jacqueline E  
; APPLICANT: Haake, Volker  
; APPLICANT: Creelman, Robert A  
; APPLICANT: Ratcliffe, Oliver  
; APPLICANT: Adam, Luc J  
; APPLICANT: Reuber, T. Lynne  
; APPLICANT: Keddie, James  
; APPLICANT: Broun, Pierre E  
; APPLICANT: Pilgrim, Marsha L  
; APPLICANT: Dubell III, Arnold T  
; APPLICANT: Pineda, Omaira  
; APPLICANT: Yu, Guo-Liang  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS

```

; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1363
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Orthologous to G1051, G1052
US-10-374-780A-1363

```

```

Query Match          92.9%; Score 39; DB 15; Length 324;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches      5; Conservative      1; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      1 GHHHHS 6
        |||||:
Db      21 GHHHHA 26

```

# RESULT 5

US-10-253-007-46

```

; Sequence 46, Application US/10253007
; Publication No. US20030088073A1
; GENERAL INFORMATION:
; APPLICANT: Benfey et al.
; TITLE OF INVENTION: Scarecrow Gene, Promoter and Uses
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: 5914-074-999
; CURRENT APPLICATION NUMBER: US/10/253,007
; CURRENT FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: US/09/186,188
; PRIOR FILING DATE: 1998-11-05
; PRIOR APPLICATION NUMBER: 08/842,445
; PRIOR FILING DATE: 1997-04-24
; PRIOR APPLICATION NUMBER: 08/638,617
; PRIOR FILING DATE: 1996-04-26

```

; NUMBER OF SEQ ID NOS: 79  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 46  
; LENGTH: 379  
; TYPE: PRT  
; ORGANISM: Plant  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(379)  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-10-253-007-46

Query Match 92.9%; Score 39; DB 14; Length 379;  
Best Local Similarity 83.3%; Pred. No. 1.4e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
|||||:  
Db 6 GHHHHT 11

RESULT 6

US-09-922-011-10

; Sequence 10, Application US/09922011  
; Publication No. US20030096331A1  
; GENERAL INFORMATION:  
; APPLICANT: CIS Biotech, Inc.  
; APPLICANT: Dambinova, Svetlana  
; TITLE OF INVENTION: Rapid multiple panel of biomarkers in laboratory blood tests for  
; TITLE OF INVENTION: TIA/stroke  
; FILE REFERENCE: 08805.105001  
; CURRENT APPLICATION NUMBER: US/09/922,011  
; CURRENT FILING DATE: 2001-08-02  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 10  
; LENGTH: 1480  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-09-922-011-10

Query Match 92.9%; Score 39; DB 10; Length 1480;  
Best Local Similarity 83.3%; Pred. No. 4.5e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
|||||:  
Db 1360 GHHHHN 1365

RESULT 7

US-09-945-901-56

; Sequence 56, Application US/09945901  
; Patent No. US20020161215A1  
; GENERAL INFORMATION:  
; APPLICANT: Daggett, Lorrie P.

```

; Ellis, Steven B.
; Liaw, Chen W.
; Lu, Chin-Chun
; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR
; SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND USES
THEREFOR
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Heller Ehrman White & McAuliffe
; STREET: 4250 Executive Square, 7th Floor
; CITY: La Jolla
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/945,901
; FILING DATE: 24-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/940,035
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US 08/052,449
; FILING DATE: 20-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-9383E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; INFORMATION FOR SEQ ID NO: 56:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1484 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 56:
US-09-945-901-56

```

```

Query Match          92.9%; Score 39; DB 9; Length 1484;
Best Local Similarity 83.3%; Pred. No. 4.5e+02;
Matches      5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 GHHHHS 6
        |||||:
Db      1360 GHHHHN 1365

```

```

RESULT 8
US-10-007-747-56
; Sequence 56, Application US/10007747
; Publication No. US20020161193A1

```

```

; GENERAL INFORMATION:
;   APPLICANT: Daggett, Lorrie P.
;             Ellis, Steven B.
;             Liaw, Chen W.
;             Lu, Chin-Chun
;   TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR
;                       SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND USES
THEREFOR
;   NUMBER OF SEQUENCES: 63
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Heller Ehrman White & McAuliffe
;     STREET: 4250 Executive Square, 7th Floor
;     CITY: La Jolla
;     STATE: CA
;     COUNTRY: USA
;     ZIP: 92037
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/10/007,747
;     FILING DATE: 07-Dec-2001
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US/09/648,797
;     FILING DATE: 28-Aug-2000
;     APPLICATION NUMBER: US/08/940,086A
;     FILING DATE: 29-SEPT-97
;     APPLICATION NUMBER: US 08/231,193
;     FILING DATE: 20-APR-1994
;     APPLICATION NUMBER: US 08/052,449
;     FILING DATE: 20-APR-1993
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Seidman, Stephanie
;     REGISTRATION NUMBER: 33,779
;     REFERENCE/DOCKET NUMBER: 24735-9383C
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (619) 450-8400
;     TELEFAX: (619) 450-8499
;   INFORMATION FOR SEQ ID NO: 56:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 1484 amino acids
;       TYPE: amino acid
;       TOPOLOGY: linear
;     MOLECULE TYPE: protein
;     SEQUENCE DESCRIPTION: SEQ ID NO: 56:
US-10-007-747-56

```

```

Query Match          92.9%;  Score 39;  DB 13;  Length 1484;
Best Local Similarity 83.3%;  Pred. No. 4.5e+02;
Matches      5;  Conservative    1;  Mismatches    0;  Indels      0;  Gaps      0;

```

```

Qy      1 GHHHHS 6
        ||||:
Db     1360 GHHHHN 1365

```

RESULT 9  
 US-10-038-937-56  
 ; Sequence 56, Application US/10038937  
 ; Publication No. US20030013866A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Daggett, Lorrie P.  
 ; Lu, Chin-Chun  
 ; TITLE OF INVENTION: HUMAN N-METHYL-D-ASPARTATE RECEPTOR  
 ; SUBUNITS, NUCLEIC ACIDS ENCODING SAME AND USES  
 THEREFOR  
 ; NUMBER OF SEQUENCES: 63  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Heller Ehrman White & McAuliffe  
 ; STREET: 4250 Executive Square, 7th Floor  
 ; CITY: La Jolla  
 ; STATE: CA  
 ; COUNTRY: U.S.A.  
 ; ZIP: 92037  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/10/038,937  
 ; FILING DATE: 18-Apr-2002  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/935,105  
 ; FILING DATE: 29-SEPT-97  
 ; APPLICATION NUMBER: US 08/231,193  
 ; FILING DATE: 20-APR-1994  
 ; APPLICATION NUMBER: US 08/052,449  
 ; FILING DATE: 20-APR-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Seidman, Stephanie  
 ; REGISTRATION NUMBER: 33,779  
 ; REFERENCE/DOCKET NUMBER: 6362-9383D  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 619-238-0999  
 ; TELEFAX: 619-238-0062  
 ; INFORMATION FOR SEQ ID NO: 56:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1484 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 56:  
 US-10-038-937-56

Query Match 92.9%; Score 39; DB 14; Length 1484;  
 Best Local Similarity 83.3%; Pred. No. 4.5e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
 |||||:



Db 1360 GHHHHN 1365

RESULT 10

US-10-146-806-2

; Sequence 2, Application US/10146806

; Publication No. US20030087371A1

; GENERAL INFORMATION:

; APPLICANT: FOLDES, Robert L.

; ADAMS, Sally-Lin

; KAMBOJ, Rajender

; DUNCAN, H. Scott

; TITLE OF INVENTION: Modulatory Proteins of Human CNS  
; Receptors

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 3000 K Street, N.W., Suite 500

; CITY: Washington, D.C.

; COUNTRY: USA

; ZIP: 20007-5109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/146,806

; FILING DATE: 17-May-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/264,578

; FILING DATE: 23-JUN-1994

; APPLICATION NUMBER: US 07/987,953

; FILING DATE: 11-DEC-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: BENT, Stephen A.

; REGISTRATION NUMBER: 29,768

; REFERENCE/DOCKET NUMBER: 16777/261/ALLE

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)672-5300

; TELEFAX: (202)672-5399

; TELEX: 904136

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1484 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-10-146-806-2

Query Match 92.9%; Score 39; DB 14; Length 1484;

Best Local Similarity 83.3%; Pred. No. 4.5e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6

Db           1360 GHHHHN 1365

RESULT 11

US-10-179-784-39

; Sequence 39, Application US/10179784  
; Publication No. US20030036647A1  
; GENERAL INFORMATION:  
; APPLICANT: Shuman, Stewart  
; APPLICANT: Sriskanda, Verl  
; TITLE OF INVENTION: Pharmacological Targeting of Bacterial DNA Ligase  
; TITLE OF INVENTION: For Treatment And Prevention of Bacterial Infections  
; FILE REFERENCE: D6468  
; CURRENT APPLICATION NUMBER: US/10/179,784  
; CURRENT FILING DATE: 2002-06-24  
; PRIOR APPLICATION NUMBER: US 60/300,727  
; PRIOR FILING DATE: 2001-06-24  
; NUMBER OF SEQ ID NOS: 41  
; SEQ ID NO 39  
; LENGTH: 6  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; NAME/KEY: CHAIN  
; OTHER INFORMATION: a histine tag  
US-10-179-784-39

Query Match                   90.5%; Score 38; DB 14; Length 6;  
Best Local Similarity   100.0%; Pred. No. 7.1e+05;  
Matches       5; Conservative   0; Mismatches   0; Indels       0; Gaps       0;

Qy           1 GHHHH 5  
              |||||  
Db           1 GHHHH 5

RESULT 12

US-09-821-984-44

; Sequence 44, Application US/09821984  
; Patent No. US20020004205A1  
; GENERAL INFORMATION:  
; APPLICANT: Consler, Thomas G.  
; APPLICANT: Iannone, Marie A.  
; APPLICANT: Gray, John G.  
; APPLICANT: Stimmel, Julia E.  
; TITLE OF INVENTION: METHOD OF INVESTIGATING FUNCTIONAL  
; TITLE OF INVENTION: MOLECULAR INTERACTIONS AND REAGENTS FOR USE THEREIN  
; FILE REFERENCE: 07083.0007U2  
; CURRENT APPLICATION NUMBER: US/09/821,984  
; CURRENT FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 60/193,826  
; PRIOR FILING DATE: 2000-03-31  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 44  
; LENGTH: 9

; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:/note =  
; OTHER INFORMATION: synthetic construct  
US-09-821-984-44

Query Match 90.5%; Score 38; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.1e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5  
|||||  
Db 4 GHHHH 8

RESULT 13

US-09-284-663A-25  
; Sequence 25, Application US/09284663A  
; Patent No. US20020012961A1  
; GENERAL INFORMATION:  
; APPLICANT: Botstein, David A.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Lawrence, David A.  
; APPLICANT: Roy, Margaret Ann  
; TITLE OF INVENTION: Fibroblast Growth Factor-19  
; FILE REFERENCE: P1219R1(e)  
; CURRENT APPLICATION NUMBER: US/09/284,663A  
; CURRENT FILING DATE: 1999-04-15  
; NUMBER OF SEQ ID NOS: 30  
; SEQ ID NO 25  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic epitope-tag.  
US-09-284-663A-25

Query Match 90.5%; Score 38; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.1e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5  
|||||  
Db 1 GHHHH 5

RESULT 14

US-09-854-280-18  
; Sequence 18, Application US/09854280  
; Patent No. US20020052027A1  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Jian  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Goddard, Audrey

```

; APPLICANT: Gurney, Austin
; APPLICANT: Li, Hanzhong
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: IL-17 HOMOLOGOUS POLYPEPTIDES AND THERAPEUTIC USES
THEREOF
; FILE REFERENCE: P1381R1C2
; CURRENT APPLICATION NUMBER: US/09/854,280
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: US 09/311,832
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/085,579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: US 60/113,621
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 26
; SEQ ID NO 18
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HIS tag
US-09-854-280-18

```

```

Query Match          90.5%; Score 38; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches      5; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

```

```

Qy      1 GHHHH 5
        |||||
Db      1 GHHHH 5

```

```

RESULT 15
US-09-854-208-18
; Sequence 18, Application US/09854208
; Patent No. US20020106743A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Jian
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin
; APPLICANT: Li, Hanzhong
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: IL-17 HOMOLOGOUS POLYPEPTIDES AND THERAPEUTIC USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: P1381-R1
; CURRENT APPLICATION NUMBER: US/09/854,208
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: US/09/311,832
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: US 60/085,579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: US 60/113,621
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 26
; SEQ ID NO 18
; LENGTH: 9

```

; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: Artificial Sequence  
; LOCATION: 1-9  
; OTHER INFORMATION: His tag  
US-09-854-208-18

Query Match 90.5%; Score 38; DB 9; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.1e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5  
|||||  
Db 1 GHHHH 5

Search completed: March 5, 2004, 16:33:44  
Job time : 3.24074 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 5, 2004, 16:15:44 ; Search time 3.98148 Seconds  
(without alignments)  
475.479 Million cell updates/sec

Title: US-10-057-890A-15  
Perfect score: 42  
Sequence: 1 GHHHHS 6  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_25:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp Vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query	%
No.	Score	Match Length DB ID Description
-----		

1	42	100.0	172	16	Q7USF2	Q7usf2 rhodopirell
2	42	100.0	353	10	Q8LFE3	Q8lfe3 arabidopsis
3	42	100.0	597	10	Q8LGQ3	Q8lgq3 oryza sativ
4	42	100.0	650	16	Q892X9	Q892x9 clostridium
5	42	100.0	692	10	Q8S1E1	Q8s1e1 oryza sativ
6	42	100.0	1056	5	Q9VI26	Q9vi26 drosophila
7	39	92.9	70	10	Q8LE23	Q8le23 arabidopsis
8	39	92.9	140	5	Q9XWP9	Q9xwp9 caenorhabdi
9	39	92.9	239	5	Q9W2R5	Q9w2r5 drosophila
10	39	92.9	278	10	Q9AR62	Q9ar62 solanum tub
11	39	92.9	278	10	Q9AR64	Q9ar64 solanum tub
12	39	92.9	307	10	Q8LR43	Q8lr43 oryza sativ
13	39	92.9	361	16	Q9ABC7	Q9abc7 caulobacter
14	39	92.9	378	10	Q9SYQ4	Q9syq4 arabidopsis
15	39	92.9	558	10	O81316	O81316 arabidopsis
16	39	92.9	782	5	Q8I4I2	Q8i4i2 caenorhabdi
17	39	92.9	1099	16	Q88EP4	Q88ep4 pseudomonas
18	39	92.9	1880	5	Q8MP27	Q8mp27 dictyosteli
19	39	92.9	1922	5	Q8I2P4	Q8i2p4 plasmodium
20	38	90.5	49	5	Q86H74	Q86h74 dictyosteli
21	38	90.5	76	5	Q25550	Q25550 naegleria f
22	38	90.5	77	5	Q20690	Q20690 caenorhabdi
23	38	90.5	77	16	Q7UQI0	Q7uqi0 rhodopirell
24	38	90.5	81	5	Q86HB7	Q86hb7 dictyosteli
25	38	90.5	83	5	Q20689	Q20689 caenorhabdi
26	38	90.5	84	5	Q86IJ6	Q86ij6 dictyosteli
27	38	90.5	89	5	Q86IJ5	Q86ij5 dictyosteli
28	38	90.5	93	5	Q86IJ4	Q86ij4 dictyosteli
29	38	90.5	99	16	Q8EN35	Q8en35 oceanobacil
30	38	90.5	100	16	Q98FY5	Q98fy5 rhizobium l
31	38	90.5	102	5	Q9VUE1	Q9vuel drosophila
32	38	90.5	102	5	Q94189	Q94189 caenorhabdi
33	38	90.5	102	16	Q83BC2	Q83bc2 coxiella bu
34	38	90.5	104	5	Q86IK1	Q86ik1 dictyosteli
35	38	90.5	109	2	Q9KI86	Q9ki86 bacillus an
36	38	90.5	109	2	Q9KI85	Q9ki85 bacillus an
37	38	90.5	112	2	Q9KI84	Q9ki84 bacillus an
38	38	90.5	112	10	Q93VD6	Q93vd6 cucumis mel
39	38	90.5	112	10	Q7XYV5	Q7xyv5 lycopersico
40	38	90.5	112	16	Q81V73	Q81v73 bacillus an
41	38	90.5	114	10	Q40165	Q40165 lycopersico
42	38	90.5	114	10	Q7XYV6	Q7xyv6 lycopersico
43	38	90.5	114	10	Q7XYV4	Q7xyv4 lycopersico
44	38	90.5	114	10	Q7XYV3	Q7xyv3 lycopersico
45	38	90.5	115	2	Q9KI83	Q9ki83 bacillus an

# ALIGNMENTS

## RESULT 1

Q7USF2

ID Q7USF2 PRELIMINARY; PRT; 172 AA.

AC Q7USF2;

DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hypothetical protein.  
 GN RB4538.  
 OS Rhodopirellula baltica.  
 OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;  
 OC Planctomycetaceae; Pirellula.  
 OX NCBI\_TaxID=117;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1;  
 RX MEDLINE=22735913; PubMed=12835416;  
 RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,  
 RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,  
 RA Schlesner H., Amann R., Reinhardt R.;  
 RT "Complete genome sequence of the marine planctomycete Pirellula sp.  
 RT strain 1.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).  
 DR EMBL; BX294140; CAD73845.1; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 172 AA; 18532 MW; 124AC22579E5B6FC CRC64;

Query Match 100.0%; Score 42; DB 16; Length 172;  
 Best Local Similarity 100.0%; Pred. No. 9.3;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHHHS 6  
 |||||  
 Db 52 GHHHHS 57

## RESULT 2

### Q8LFE3

ID Q8LFE3 PRELIMINARY; PRT; 353 AA.  
 AC Q8LFE3;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Zinc finger protein, putative.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,  
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;  
 RT "Full-length messenger RNA sequences greatly improve genome  
 RT annotation.";  
 RL Genome Biol. 0:0-0(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,  
 RA Feldmann K.;  
 RT "Full-Length cDNA from Arabidopsis thaliana.";  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AY084898; AAM61461.1; -.  
 DR GO; GO:0003677; F:DNA binding; IEA.



DR InterPro; IPR003851; Znf\_Dof.  
 DR Pfam; PF02701; zf-Dof; 1.  
 DR PROSITE; PS01361; ZF\_DOF\_1; 1.  
 DR PROSITE; PS50884; ZF\_DOF\_2; 1.  
 SQ SEQUENCE 353 AA; 37960 MW; C97EBE7B09A2E6FA CRC64;

Query Match 100.0%; Score 42; DB 10; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
 |||||  
 Db 178 GHHHHS 183

# RESULT 3

## Q8LGQ3

ID Q8LGQ3 PRELIMINARY; PRT; 597 AA.  
 AC Q8LGQ3;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Ovule development aintegumenta-like protein BNM3.  
 GN BNM3.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bi X.-Z.;  
 RT "Cloning and identification of two ovule development proteins,  
 RT aintegumenta-like protein in rice (Oryza sativa).";  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AY062180; AAL47205.1; -.  
 DR Gramene; Q8LGQ3; -.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR001471; TF\_ERF.  
 DR Pfam; PF00847; AP2-domain; 2.  
 DR PRINTS; PR00367; ETHRSPELEMNT.  
 DR ProDom; PD001423; TF\_ERF; 2.  
 DR SMART; SM00380; AP2; 2.  
 SQ SEQUENCE 597 AA; 62198 MW; F856EBC99BADE25B CRC64;

Query Match 100.0%; Score 42; DB 10; Length 597;  
 Best Local Similarity 100.0%; Pred. No. 30;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
 |||||  
 Db 398 GHHHHS 403

# RESULT 4

Q892X9

ID Q892X9 PRELIMINARY; PRT; 650 AA.  
AC Q892X9;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Zinc-transporting ATPase (EC 3.6.1.-).  
GN CTC01955.  
OS Clostridium tetani.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1513;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Massachusetts / E88;  
RX MEDLINE=22457253; PubMed=12552129;  
RA Brueggemann H., Baeumer S., Fricke W.F., Wiezer A., Liesegang H.,  
RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,  
RA Gottschalk G.;  
RT "The genome sequence of Clostridium tetani, the causative agent of  
RT tetanus disease.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).  
DR EMBL; AE015942; AAC036463.1; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0015662; F:ATPase activity, coupled to transmembrane m. . .; IEA.  
DR GO; GO:0016787; F:hydrolase activity; IEA.  
DR GO; GO:0006812; P:cation transport; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR InterPro; IPR001757; ATPase\_E1-E2.  
DR InterPro; IPR008250; E1-E2\_ATPase\_reg.  
DR InterPro; IPR005834; Hydrolase.  
DR Pfam; PF00122; E1-E2\_ATPase; 1.  
DR Pfam; PF00702; Hydrolase; 1.  
DR PRINTS; PR00119; CATATPASE.  
DR PROSITE; PS00154; ATPASE\_E1\_E2; 1.  
KW Hydrolase; Complete proteome.  
SQ SEQUENCE 650 AA; 71407 MW; F1F7800A09EE0793 CRC64;

Query Match 100.0%; Score 42; DB 16; Length 650;  
Best Local Similarity 100.0%; Pred. No. 33;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
      |||||  
Db 13 GHHHHS 18

RESULT 5

Q8S1E1

ID Q8S1E1 PRELIMINARY; PRT; 692 AA.  
AC Q8S1E1;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Putative ovule development protein aintegumenta-like protein.  
GN P0035F12.3.

OS *Oryza sativa* (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; *Oryza*.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC  
 RT clone:P0035F12.";  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AP003313; BAB89946.1; -.  
 DR Gramene; Q8S1E1; -.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR001471; TF\_ERF.  
 DR Pfam; PF00847; AP2-domain; 2.  
 DR PRINTS; PR00367; ETHRSPELEMNT.  
 DR ProDom; PD001423; TF\_ERF; 2.  
 DR SMART; SM00380; AP2; 2.  
 SQ SEQUENCE 692 AA; 71515 MW; 4D5A0B49ED8772AF CRC64;

Query Match 100.0%; Score 42; DB 10; Length 692;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
 |||||  
 Db 493 GHHHHS 498

# RESULT 6

## Q9VI26

ID Q9VI26 PRELIMINARY; PRT; 1056 AA.  
 AC Q9VI26;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE CG15186 protein.  
 GN CG15186.  
 OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; *Drosophila*.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkeley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AE003674; AAF54118.1; -.  
 DR FlyBase; FBgn0037448; CG15186.  
 SQ SEQUENCE 1056 AA; 113358 MW; EC8BC31402D7FE52 CRC64;

Query Match 100.0%; Score 42; DB 5; Length 1056;  
 Best Local Similarity 100.0%; Pred. No. 52;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHHHS 6  
 |||||  
 Db 969 GHHHHS 974

# RESULT 7

Q8LE23

ID Q8LE23 PRELIMINARY; PRT; 70 AA.

AC Q8LE23;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE Hypothetical protein.

OS Arabidopsis thaliana (Mouse-ear cress).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,  
 RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;  
 RT "Full-length messenger RNA sequences greatly improve genome  
 RT annotation.";  
 RL Genome Biol. 0:0-0(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,  
 RA Feldmann K.;  
 RT "Full-Length cDNA from Arabidopsis thaliana.";  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AY085662; AAM67306.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 70 AA; 7270 MW; 10C0764E0986E031 CRC64;

Query Match 92.9%; Score 39; DB 10; Length 70;  
 Best Local Similarity 83.3%; Pred. No. 12;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
 |||||:  
 Db 29 GHHHHA 34

#### RESULT 8

##### Q9XWP9

ID Q9XWP9 PRELIMINARY; PRT; 140 AA.  
 AC Q9XWP9;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Y51A2A.6 protein.  
 GN Y51A2A.6.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA McMurray A.A.;  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology.";  
 RL Science 282:2012-2018(1998).  
 DR EMBL; AL032635; CAA21601.1; -.  
 DR PIR; T27059; T27059.  
 DR WormPep; Y51A2A.6; CE20277.  
 SQ SEQUENCE 140 AA; 15367 MW; 1F57C3547BE07568 CRC64;

Query Match 92.9%; Score 39; DB 5; Length 140;

Best Local Similarity 83.3%; Pred. No. 23;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHHHS 6  
|||||:  
Db 53 GHHHHN 58

RESULT 9

Q9W2R5

ID Q9W2R5 PRELIMINARY; PRT; 239 AA.

AC Q9W2R5;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE CG15225 protein.

GN CG15225.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkeley;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AE003452; AAF46625.1; -.  
 DR FlyBase; FBgn0034551; CG15225.  
 SQ SEQUENCE 239 AA; 26175 MW; 81EEE85DD2FC5FB7 CRC64;

Query Match 92.9%; Score 39; DB 5; Length 239;  
 Best Local Similarity 83.3%; Pred. No. 39;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
 |||||:  
 Db 162 GHHHHT 167

# RESULT 10

Q9AR62

ID Q9AR62 PRELIMINARY; PRT; 278 AA.  
 AC Q9AR62;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Urease accessory protein G.  
 GN UREG.  
 OS Solanum tuberosum (Potato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC lamiids; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4113;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Desiree; TISSUE=Leaf;  
 RX MEDLINE=21183143; PubMed=11289508;  
 RA Witte C.P., Isidore E., Tiller S.A., Davies H.V., Taylor M.A.;  
 RT "Functional characterisation of urease accessory protein G (ureG) in  
 RT potato.";  
 RL Plant Mol. Biol. 45:169-179(2001).  
 DR EMBL; AJ272525; CAC33002.1; -.  
 DR GO; GO:0046872; F:metal ion binding; IEA.  
 DR GO; GO:0016151; F:nickel ion binding; IEA.  
 DR GO; GO:0000166; F:nucleotide binding; IEA.  
 DR GO; GO:0006461; P:protein complex assembly; IEA.  
 DR InterPro; IPR002894; HypB\_UreG.  
 DR InterPro; IPR004400; UreG.  
 DR Pfam; PF01495; HypB\_UreG; 1.  
 DR TIGRFAMs; TIGR00101; ureG; 1.  
 SQ SEQUENCE 278 AA; 30373 MW; E0999F79ED8BA478 CRC64;

Query Match 92.9%; Score 39; DB 10; Length 278;  
 Best Local Similarity 83.3%; Pred. No. 45;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
|||||:  
Db 14 GHHHHN 19

RESULT 11

Q9AR64

ID Q9AR64 PRELIMINARY; PRT; 278 AA.  
AC Q9AR64;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Urease accessory protein G.  
GN UREG.  
OS Solanum tuberosum (Potato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC lamiids; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4113;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Record; TISSUE=Stolon;  
RX MEDLINE=21183143; PubMed=11289508;  
RA Witte C.P., Isidore E., Tiller S.A., Davies H.V., Taylor M.A.;  
RT "Functional characterisation of urease accessory protein G (ureG) in  
RT potato.";  
RL Plant Mol. Biol. 45:169-179(2001).  
DR EMBL; AJ272523; CAC33000.1; -.  
DR GO; GO:0046872; F:metal ion binding; IEA.  
DR GO; GO:0016151; F:nickel ion binding; IEA.  
DR GO; GO:0000166; F:nucleotide binding; IEA.  
DR GO; GO:0006461; P:protein complex assembly; IEA.  
DR InterPro; IPR002894; HypB\_UreG.  
DR InterPro; IPR004400; UreG.  
DR Pfam; PF01495; HypB\_UreG; 1.  
DR TIGRFAMs; TIGR00101; ureG; 1.  
SQ SEQUENCE 278 AA; 30357 MW; 51530F68FD8E2918 CRC64;

Query Match 92.9%; Score 39; DB 10; Length 278;  
Best Local Similarity 83.3%; Pred. No. 45;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
|||||:  
Db 14 GHHHHN 19

RESULT 12

Q8LR43

ID Q8LR43 PRELIMINARY; PRT; 307 AA.  
AC Q8LR43;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE P0512C01.32 protein.  
GN P0512C01.32.  
OS Oryza sativa (japonica cultivar-group).



OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC  
 RT clone:P0512C01.";  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AP003274; BAB92377.1; -.  
 DR Gramene; Q8LR43; -.  
 DR InterPro; IPR005333; TCP.  
 DR Pfam; PF03634; TCP; 1.  
 SQ SEQUENCE 307 AA; 32004 MW; 776182ACA7178987 CRC64;

Query Match 92.9%; Score 39; DB 10; Length 307;  
 Best Local Similarity 83.3%; Pred. No. 49;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
 ||||:  
 Db 144 GHHHHA 149

# RESULT 13

## Q9ABC7

ID Q9ABC7 PRELIMINARY; PRT; 361 AA.  
 AC Q9ABC7;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Cation efflux family protein.  
 GN CC0303.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;  
 OC Caulobacteraceae; Caulobacter.  
 OX NCBI\_TaxID=155892;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 19089 / CB15;  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
 RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of Caulobacter crescentus."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL; AE005704; AAK22290.1; -.  
 DR PIR; F87286; F87286.  
 DR TIGR; CC0303; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0008324; F:cation transporter activity; IEA.

DR GO; GO:0006812; P:cation transport; IEA.  
DR InterPro; IPR002524; Cation\_efflux.  
DR Pfam; PF01545; Cation\_efflux; 1.  
DR TIGRFAMs; TIGR01297; CDF; 1.  
KW Complete proteome.  
SQ SEQUENCE 361 AA; 38180 MW; 1A4F7F0A7C62EEB0 CRC64;

Query Match 92.9%; Score 39; DB 16; Length 361;  
Best Local Similarity 83.3%; Pred. No. 57;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHHHS 6  
|||  
Db 64 GHHHHA 69

#### RESULT 14

Q9SYQ4

ID Q9SYQ4 PRELIMINARY; PRT; 378 AA.  
AC Q9SYQ4;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Scarecrow-like 6 (Fragment).  
GN SCL6.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99272994; PubMed=10341448;  
RA Pysh L.D., Wysocka-Diller J.W., Camilleri C., Bouchez D., Benfey P.N.;  
RT "The GRAS gene family in Arabidopsis: sequence characterization and  
RT basic expression analysis of the SCARECROW-LIKE genes.";  
RL Plant J. 18:111-119(1999).  
DR EMBL; AF036303; AAD24406.1; -.  
DR PIR; T51237; T51237.  
DR InterPro; IPR005202; GRAS.  
DR Pfam; PF03514; GRAS; 1.  
FT NON\_TER 1 1  
SQ SEQUENCE 378 AA; 42321 MW; CA7FDC7C09B2CB21 CRC64;

Query Match 92.9%; Score 39; DB 10; Length 378;  
Best Local Similarity 83.3%; Pred. No. 60;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHHS 6  
|||  
Db 6 GHHHHT 11

#### RESULT 15

O81316

ID O81316 PRELIMINARY; PRT; 558 AA.  
AC O81316;

DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE F6N15.20 protein (SCARECROW-like 6) (SCL6) (AT4g00150/F6N15\_20).  
 GN F6N15.20 OR AT4G00150.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA WASHU;  
 RT "The A. thaliana Genome Sequencing Project.";  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA Ryan E., Edwards J., Pape K.;  
 RT "The sequence of A. thaliana F6N15.";  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RA Waterston R.;  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Wilson R., Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,  
 RA Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Banh J., Bowser L.,  
 RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,  
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,  
 RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,  
 RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
 RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,  
 RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
 RT "Arabidopsis cDNA clones.";  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RA Kim C.J., Chen H., Cheuk R., Shinn P., Banh J., Bowser L.,  
 RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,  
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,  
 RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,  
 RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
 RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,  
 RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
 RT "Arabidopsis ORF clones.";

RL Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF069299; AAC19296.1; -.  
DR EMBL; AL161471; CAB80773.1; -.  
DR EMBL; AF462831; AAL58919.1; -.  
DR EMBL; AY133537; AAM91367.1; -.  
DR PIR; T01343; T01343.  
DR InterPro; IPR005202; GRAS.  
DR Pfam; PF03514; GRAS; 1.  
SQ SEQUENCE 558 AA; 61167 MW; FA07D1A11B053910 CRC64;

Query Match 92.9%; Score 39; DB 10; Length 558;  
Best Local Similarity 83.3%; Pred. No. 87;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHHHS 6  
|||||:  
Db 186 GHHHHT 191

Search completed: March 5, 2004, 16:27:31  
Job time : 3.98148 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 5, 2004, 16:15:14 ; Search time 0.814815 Seconds  
(without alignments)  
383.426 Million cell updates/sec

Title: US-10-057-890A-15  
Perfect score: 42  
Sequence: 1 GHHHHS 6

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	39	92.9	1484	1	NME2_HUMAN	Q13224 homo sapien
2	38	90.5	59	1	HPN_HELPY	Q48251 helicobacte
3	38	90.5	107	1	HSP2_MOUSE	P07978 mus musculu
4	38	90.5	114	1	ASR2_LYCES	P37219 lycopersico
5	38	90.5	117	1	HIA1_DICDI	P13231 dictyosteli
6	38	90.5	117	1	HIA2_DICDI	P42526 dictyosteli
7	38	90.5	131	1	INL3_PIG	P51461 sus scrofa
8	38	90.5	161	1	UREE_PROMI	P17090 proteus mir
9	38	90.5	274	1	YOHM_ECOLI	P76425 escherichia
10	38	90.5	298	1	MOX2_XENLA	P39021 xenopus lae
11	38	90.5	299	1	HYPB_RHILV	P28155 rhizobium l
12	38	90.5	302	1	HYPB_BRAJA	Q45257 bradyrhizob
13	38	90.5	303	1	MOX2_HUMAN	P50222 homo sapien
14	38	90.5	303	1	MOX2_MOUSE	P32443 mus musculu
15	38	90.5	303	1	MOX2_RAT	P39020 rattus norv
16	38	90.5	305	1	HYPB_AZOCH	Q43949 azotobacter
17	38	90.5	323	1	OTX1_BRARE	Q91994 brachydanio

18	38	90.5	338	1	IAR1_ARATH	Q9m647 arabidopsis
19	38	90.5	351	1	HRPX_PLALO	P04929 plasmodium
20	38	90.5	354	1	OTX1_HUMAN	P32242 homo sapien
21	38	90.5	355	1	OTX1_MOUSE	P80205 mus musculu
22	38	90.5	355	1	OTX1_RAT	Q63410 rattus norv
23	38	90.5	369	1	MAF_RAT	P54844 rattus norv
24	38	90.5	370	1	MAF_MOUSE	P54843 mus musculu
25	38	90.5	377	1	CAH1_CHLRE	P20507 chlamydomon
26	38	90.5	380	1	CAH2_CHLRE	P24258 chlamydomon
27	38	90.5	403	1	MAF_HUMAN	O75444 homo sapien
28	38	90.5	411	1	NORV_ECO57	Q8x852 escherichia
29	38	90.5	414	1	TYY1_HUMAN	P25490 homo sapien
30	38	90.5	414	1	TYY1_MOUSE	Q00899 mus musculu
31	38	90.5	420	1	YBE1_SCHPO	O42980 schizosacch
32	38	90.5	443	1	ZIC1_XENLA	O73689 xenopus lae
33	38	90.5	447	1	ZIC1_HUMAN	Q15915 homo sapien
34	38	90.5	447	1	ZIC1_MOUSE	P46684 mus musculu
35	38	90.5	449	1	CSUP_DROME	Q9v3a4 drosophila
36	38	90.5	466	1	ZIC3_MOUSE	Q62521 mus musculu
37	38	90.5	467	1	ZIC3_HUMAN	O60481 homo sapien
38	38	90.5	479	1	NORV_ECOL6	P59404 escherichia
39	38	90.5	479	1	NORV_ECOLI	Q46877 escherichia
40	38	90.5	479	1	NORV_SALTI	Q8z4c5 salmonella
41	38	90.5	479	1	NORV_SALTY	Q8zmj7 salmonella
42	38	90.5	479	1	NORV_SHIFL	P59405 shigella fl
43	38	90.5	494	1	NORV_VIBVU	Q8d4f8 vibrio vuln
44	38	90.5	496	1	BAF1_KLUMA	P33293 kluyveromyc
45	38	90.5	503	1	YKR5_YEAST	P34240 saccharomyc

# ALIGNMENTS

## RESULT 1

### NME2\_HUMAN

ID NME2\_HUMAN STANDARD; PRT; 1484 AA.  
AC Q13224; Q12919; Q13220; Q13225; Q9UM56;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Glutamate [NMDA] receptor subunit epsilon 2 precursor (N-methyl D-  
DE aspartate receptor subtype 2B) (NR2B) (NMDAR2B) (N-methyl-D-aspartate  
DE receptor subunit 3) (NR3) (hNR3).  
GN GRIN2B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND VARIANT ASN-407.  
RC TISSUE=Fetal brain;  
RX MEDLINE=95092783; PubMed=7999784;  
RA Adams S.L., Foldes R.L., Kamboj R.K.;  
RT "Human N-methyl-D-aspartate receptor modulatory subunit hNR3: cloning  
RT and sequencing of the cDNA and primary structure of the protein.";  
RL Biochim. Biophys. Acta 1260:105-108(1995).  
RN [2]

RP SEQUENCE FROM N.A.  
 RC TISSUE=Fetal brain;  
 RX MEDLINE=96312186; PubMed=8768735;  
 RA Hess S.D., Daggett L.P., Crona J., Deal C., Lu C.-C., Urrutia A.,  
 RA Chavez-Noriega L., Ellis S.B., Johnson E.C., Velicelebi G.;  
 RT "Cloning and functional characterization of human heteromeric N-  
 RT methyl-D-aspartate receptors.";  
 RL J. Pharmacol. Exp. Ther. 278:808-816(1996).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Mandich P., Schito A.M., Pizzuti A., Ratti A.;  
 RT "Cloning of GRIN2B human subunit.";  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DDBJ databases.  
 RN [4]  
 RP SEQUENCE OF 1-294 AND 661-1089 FROM N.A.  
 RX MEDLINE=95048375; PubMed=7959773;  
 RA Mandich P., Schito A.M., Bellone E., Antonacci R., Finelli P.,  
 RA Rocchi M., Ajmar F.;  
 RT "Mapping of the human NMDAR2B receptor subunit gene (GRIN2B) to  
 RT chromosome 12p12.";  
 RL Genomics 22:216-218(1994).  
 RN [5]  
 RP TISSUE SPECIFICITY.  
 RX MEDLINE=98140597; PubMed=9547169;  
 RA Schito A.M., Pizzuti A., Di Maria E., Schenone A., Ratti A.,  
 RA Defferrari R., Bellone E., Mancardi G.L., Ajmar F., Mandich P.;  
 RT "mRNA distribution in adult human brain of GRIN2B, a N-methyl-D-  
 RT aspartate (NMDA) receptor subunit.";  
 RL Neurosci. Lett. 239:49-53(1997).  
 CC -!- FUNCTION: NMDA receptor subtype of glutamate-gated ion channels  
 CC with high calcium permeability and voltage-dependent sensitivity  
 CC to magnesium. Mediated by glycine.  
 CC -!- SUBUNIT: Heterodimer of an epsilon subunit and a zeta subunit.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- TISSUE SPECIFICITY: Primarily found in the fronto-parieto-temporal  
 CC cortex and hippocampus pyramidal cells, lower expression in the  
 CC basal ganglia.  
 CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U90278; AAB49993.1; -.  
 DR EMBL; U88963; AAD00659.1; -.  
 DR EMBL; U11287; AAB60368.1; -.  
 DR EMBL; U28861; AAA69919.1; -.  
 DR EMBL; U28862; AAA69920.1; -.  
 DR EMBL; U28758; AAA74930.1; -.  
 DR PIR; I39066; I39066.  
 DR PIR; S52086; S52086.  
 DR HSSP; P19491; 1GR2.  
 DR Genew; HGNC:4586; GRIN2B.

DR MIM; 138252; -.  
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.  
 DR GO; GO:0004972; F:N-methyl-D-aspartate selective glutamate re. . .; TAS.  
 DR GO; GO:0007215; P:glutamate signaling pathway; TAS.  
 DR GO; GO:0007611; P:learning and/or memory; TAS.  
 DR GO; GO:0007268; P:synaptic transmission; TAS.  
 DR GO; GO:0006810; P:transport; TAS.  
 DR InterPro; IPR001320; Ion\_glu\_receptor.  
 DR InterPro; IPR001508; NMDA\_receptor.  
 DR InterPro; IPR001311; SBP/glu\_receptor.  
 DR Pfam; PF00060; lig\_chan; 1.  
 DR PRINTS; PR00177; NMDARECEPTOR.  
 DR SMART; SM00079; PBPe; 1.  
 KW Receptor; Signal; Transmembrane; Postsynaptic membrane; Calcium;  
 KW Ionic channel; Magnesium; Glycoprotein; Polymorphism; Phosphorylation.  
 FT SIGNAL 1 26 POTENTIAL.  
 FT CHAIN 27 1484 GLUTAMATE [NMDA] RECEPTOR SUBUNIT  
 FT EPSILON 2.  
 FT DOMAIN 27 557 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 558 578 1 (POTENTIAL).  
 FT DOMAIN 579 599 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 600 620 2 (POTENTIAL).  
 FT DOMAIN 621 634 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 635 655 3 (POTENTIAL).  
 FT DOMAIN 656 817 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 818 838 4 (POTENTIAL).  
 FT DOMAIN 839 1484 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 984 989 POLY-HIS.  
 FT DOMAIN 1361 1364 POLY-HIS.  
 FT SITE 615 615 FUNCTIONAL DETERMINANT OF NMDA  
 FT RECEPTORS (BY SIMILARITY).  
 FT MOD\_RES 1474 1474 PHOSPHORYLATION (BY SIMILARITY).  
 FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 341 341 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 348 348 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 491 491 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 892 892 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 910 910 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1175 1175 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1200 1200 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1224 1224 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1275 1275 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1352 1352 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1450 1450 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1466 1466 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARIANT 407 407 S -> N.  
 FT /FTid=VAR\_011317.  
 FT CONFLICT 434 434 V -> A (IN REF. 3).  
 FT CONFLICT 745 745 G -> A (IN REF. 4).  
 FT CONFLICT 773 773 K -> N (IN REF. 4).  
 FT CONFLICT 796 796 W -> C (IN REF. 4).  
 FT CONFLICT 888 888 T -> P (IN REF. 4).  
 FT CONFLICT 902 902 L -> V (IN REF. 4).  
 FT CONFLICT 920 921 SA -> RP (IN REF. 1).  
 FT CONFLICT 958 958 L -> S (IN REF. 4).



FT CONFLICT 980 982 VYQ -> DHY (IN REF. 4).  
 FT CONFLICT 1056 1056 I -> M (IN REF. 4).  
 FT CONFLICT 1167 1167 V -> I (IN REF. 2).  
 SQ SEQUENCE 1484 AA; 166366 MW; 40AEB12BE6E50CEF CRC64;

Query Match 92.9%; Score 39; DB 1; Length 1484;  
 Best Local Similarity 83.3%; Pred. No. 56;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHHHS 6  
 |||||:  
 Db 1360 GHHHHN 1365

## RESULT 2

### HPN\_HELPY

ID HPN\_HELPY STANDARD; PRT; 59 AA.  
 AC Q48251;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Histidine-rich, metal binding polypeptide.  
 GN HPN OR HP1427 OR JHP1320.  
 OS Helicobacter pylori (Campylobacter pylori), and  
 OS Helicobacter pylori J99 (Campylobacter pylori J99).  
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
 OC Helicobacteraceae; Helicobacter.  
 OX NCBI\_TaxID=210, 85963;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-18 AND 47-59.  
 RC STRAIN=LEU;  
 RX MEDLINE=95310028; PubMed=7790085;  
 RA Gilbert J.V., Ramakrishna J., Sunderman F.W. Jr., Wright A.,  
 RA Plaut A.G.;  
 RT "Protein Hpn: cloning and characterization of a histidine-rich metal-  
 RT binding polypeptide in Helicobacter pylori and Helicobacter  
 RT mustelae.";  
 RL Infect. Immun. 63:2682-2688(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=26695 / ATCC 700392;  
 RX MEDLINE=97394467; PubMed=9252185;  
 RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,  
 RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,  
 RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,  
 RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,  
 RA McKenney K., FitzGerald L.M., Lee N., Adams M.D., Hickey E.K.,  
 RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,  
 RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,  
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,  
 RA Venter J.C.;  
 RT "The complete genome sequence of the gastric pathogen Helicobacter  
 RT pylori.";  
 RL Nature 388:539-547(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=J99;

RX MEDLINE=99120557; PubMed=9923682;  
 RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,  
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,  
 RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,  
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,  
 RA Trust T.J.;  
 RT "Genomic sequence comparison of two unrelated isolates of the human  
 RT gastric pathogen *Helicobacter pylori*.";  
 RL Nature 397:176-180(1999).  
 CC -!- FUNCTION: Strongly binds nickel and zinc. Binds other metals less  
 CC strongly: Co(2+) > Cu(2+) > Cd(2+) > Mn(2+). May act to increase,  
 CC or at least to preserve, urease activity. Exact function is still  
 CC unknown.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U26361; AAA85859.1; -.  
 DR EMBL; AE000643; AAD08471.1; -.  
 DR EMBL; AE001555; AAD06898.1; -.  
 DR PIR; C64698; C64698.  
 DR TIGR; HP1427; -.  
 KW Metal-binding; Zinc; Nickel; Repeat; Complete proteome.  
 FT INIT\_MET 0 0  
 FT DOMAIN 10 23 POLY-HIS.  
 FT DOMAIN 27 32 POLY-HIS.  
 FT DOMAIN 37 54 2 X 5 AA REPEATS OF E-E-G-C-C.  
 FT REPEAT 37 41 1.  
 FT REPEAT 50 54 2.  
 SQ SEQUENCE 59 AA; 6946 MW; C3AEE3F602EC973C CRC64;

Query Match 90.5%; Score 38; DB 1; Length 59;  
 Best Local Similarity 100.0%; Pred. No. 3.1;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHHH 5  
 |||||  
 Db 9 GHHHH 13

### RESULT 3

#### HSP2\_MOUSE

ID HSP2\_MOUSE STANDARD; PRT; 107 AA.  
 AC P07978;  
 DT 01-AUG-1988 (Rel. 08, Created)  
 DT 01-AUG-1988 (Rel. 08, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Sperm histone P2 precursor (Protamine MP2).  
 GN PRM2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87257931; PubMed=3600661;  
 RA Yelick P.C., Balhorn R., Johnson P.A., Corzett M., Mazrimas J.A.,  
 RA Kleene K.C., Hecht N.B.;  
 RT "Mouse protamine 2 is synthesized as a precursor whereas mouse  
 RT protamine 1 is not.";  
 RL Mol. Cell. Biol. 7:2173-2179(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88193085; PubMed=3358932;  
 RA Johnson P.A., Pschon J.J., Yelick P.C., Palmiter R.D., Hecht N.B.;  
 RT "Sequence homologies in the mouse protamine 1 and 2 genes.";  
 RL Biochim. Biophys. Acta 950:45-53(1988).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88181903; PubMed=3445973;  
 RA Hecht N.B.;  
 RT "Gene expression during spermatogenesis.";  
 RL Ann. N.Y. Acad. Sci. 513:90-101(1987).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C129;  
 RA Schlueter G., Engel W.;  
 RL Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.  
 RN [5]  
 RP SEQUENCE OF 45-107.  
 RX MEDLINE=88294032; PubMed=3401454;  
 RA Bellve A.R., McKay D.J., Renaux B.S., Dixon G.H.;  
 RT "Purification and characterization of mouse protamines P1 and P2.  
 RT Amino acid sequence of P2.";  
 RL Biochemistry 27:2890-2897(1988).  
 RN [6]  
 RP PROCESSING.  
 RX MEDLINE=91160549; PubMed=2001695;  
 RA Elsevier S.M., Noiran J., Carre-Eusebe D.;  
 RT "Processing of the precursor of protamine P2 in mouse. Identification  
 RT of intermediates by their insolubility in the presence of sodium  
 RT dodecyl sulfate.";  
 RL Eur. J. Biochem. 196:167-175(1991).  
 RN [7]  
 RP PROCESSING.  
 RX MEDLINE=91307542; PubMed=1854346;  
 RA Carre-Eusebe D., Lederer F., Le K.H.D., Elsevier S.M.;  
 RT "Processing of the precursor of protamine P2 in mouse. Peptide  
 RT mapping and N-terminal sequence analysis of intermediates.";  
 RL Biochem. J. 277:39-45(1991).  
 CC -!- FUNCTION: Protamines substitute for histones in the chromatin of  
 CC sperm during the haploid phase of spermatogenesis. They compact  
 CC sperm DNA into a highly condensed, stable and inactive complex.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- TISSUE SPECIFICITY: Testis.  
 CC -!- SIMILARITY: Belongs to the protamine P2 family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; M16456; AAA39981.1; -.  
DR EMBL; X07626; CAA30473.1; -.  
DR EMBL; X14004; CAA32170.1; -.  
DR EMBL; M27501; AAA39986.1; -.  
DR EMBL; Z47352; CAA87411.1; -.  
DR PIR; A27809; A29995.  
DR MGD; MGI:97766; Prm2.  
DR InterPro; IPR000492; Protamine\_P2.  
DR Pfam; PF00841; protamine\_P2; 1.  
KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;  
KW Testis; DNA condensation; Nuclear protein.  
FT PROPEP 1 44  
FT CHAIN 45 107 SPERM HISTONE P2.  
FT CHAIN 2 107 PP2-A.  
FT CHAIN 12 107 PP2-C.  
FT CHAIN 21 107 PP2-D.  
SQ SEQUENCE 107 AA; 13638 MW; 66F6C3776D2DC09E CRC64;

Query Match 90.5%; Score 38; DB 1; Length 107;  
Best Local Similarity 100.0%; Pred. No. 5.6;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5  
|||  
Db 46 GHHHH 50

#### RESULT 4

##### ASR2\_LYCES

ID ASR2\_LYCES STANDARD; PRT; 114 AA.  
AC P37219;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Absciscic stress ripening protein 2.  
GN ASR2.  
OS Lycopersicon esculentum (Tomato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
OC lamiids; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4081;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Ailsa Craig;  
RX MEDLINE=95148753; PubMed=7846175;  
RA Amitai-Zeigerson H., Scolnik P.A., Bar-Zvi D.;  
RT "Genomic nucleotide sequence of tomato Asr2, a second member of the  
RT stress/ripening-induced Asr1 gene family."  
RL Plant Physiol. 106:1699-1700(1994).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; X74907; CAA52873.1; -.  
DR PIR; S37150; S37150.  
DR InterPro; IPR003496; ABA\_WDS.  
DR Pfam; PF02496; ABA\_WDS; 1.  
FT DOMAIN 6 11 POLY-HIS.  
FT DOMAIN 108 113 POLY-HIS.  
SQ SEQUENCE 114 AA; 13020 MW; AE12FBBCD3631248 CRC64;

Query Match 90.5%; Score 38; DB 1; Length 114;  
Best Local Similarity 100.0%; Pred. No. 6;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHHH 5  
|||  
Db 107 GHHHH 111

#### RESULT 5

##### HIA1\_DICDI

ID HIA1\_DICDI STANDARD; PRT; 117 AA.  
AC P13231;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Hisactophilin 1 (Histidine-rich actin-binding protein 1) (HS I).  
GN HATA OR ABPH.  
OS Dictyostelium discoideum (Slime mold).  
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.  
OX NCBI\_TaxID=44689;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89123382; PubMed=2914932;  
RA Scheel J., Ziegelbauer K., Kupke T., Humbel B.M., Noegel A.A.,  
RA Gerisch G., Schleicher M.;  
RT "Hisactophilin, a histidine-rich actin-binding protein from  
RT Dictyostelium discoideum.";  
RL J. Biol. Chem. 264:2832-2839(1989).  
RN [2]  
RP PARTIAL SEQUENCE, AND POST-TRANSLATIONAL MODIFICATIONS.  
RC STRAIN=AX2, and AX3;  
RX MEDLINE=95122497; PubMed=7822284;  
RA Hanakam F., Eckerskorn C., Lottspeich F., Mueller-Taubenberger A.,  
RA Schaefer W., Gerisch G.;  
RT "The pH-sensitive actin-binding protein hisactophilin of  
RT Dictyostelium exists in two isoforms which both are myristoylated and  
RT distributed between plasma membrane and cytoplasm.";  
RL J. Biol. Chem. 270:596-602(1995).  
RN [3]  
RP STRUCTURE BY NMR.  
RX MEDLINE=93063300; PubMed=1436061;

RA Habazetti J., Gondol D., Wiltschek R., Otlewski J., Schleicher M.,  
 RA Holak T.A.;  
 RT "Structure of hisactophilin is similar to interleukin-1 beta and  
 RT fibroblast growth factor.";  
 RL Nature 359:855-858(1992).  
 CC -!- FUNCTION: May act as an intracellular pH sensor that links  
 CC chemotactic signals to responses in the microfilament system of  
 CC the cells by nucleating actin polymerization or stabilizing the  
 CC filaments.  
 CC -!- SUBUNIT: Homodimer or heterodimer of hatA and hatB, linked by a  
 CC disulfide bond.  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic or associated with inner surface  
 CC of plasma membrane.  
 CC -!- PTM: Phosphorylated.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; J04472; AAA33218.1; -.  
 DR PIR; A31429; A31429.  
 DR PDB; 1HCD; 15-OCT-94.  
 DR PDB; 1HCE; 15-OCT-94.  
 DR DictyBase; DDB0002027; hatA.  
 DR InterPro; IPR008999; Actin\_crosslink.  
 KW Actin-binding; Repeat; Myristate; Multigene family; Phosphorylation;  
 KW 3D-structure; Lipoprotein.

FT	INIT MET	0	0	
FT	LIPID	1	1	N-myristoyl glycine.
FT	DOMAIN	7	108	CONTAINS SEVERAL HHXH REPEATS.
FT	DOMAIN	33	85	2 X 13 AA APPROXIMATE REPEATS.
FT	REPEAT	33	45	1.
FT	REPEAT	73	85	2.
FT	STRAND	1	6	
FT	TURN	9	10	
FT	STRAND	12	16	
FT	TURN	17	18	
FT	STRAND	19	23	
FT	STRAND	33	38	
FT	TURN	39	40	
FT	STRAND	41	46	
FT	TURN	47	49	
FT	STRAND	50	54	
FT	TURN	57	58	
FT	STRAND	60	63	
FT	STRAND	73	78	
FT	TURN	79	80	
FT	STRAND	81	85	
FT	HELIX	87	89	
FT	STRAND	91	95	
FT	TURN	96	98	
FT	STRAND	99	103	
FT	TURN	109	110	

FT STRAND 112 115  
SQ SEQUENCE 117 AA; 13325 MW; E5B43F1F7B5D63FD CRC64;

Query Match 90.5%; Score 38; DB 1; Length 117;  
Best Local Similarity 100.0%; Pred. No. 6.1;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GHHHH 5  
|||  
Db 86 GHHHH 90

# RESULT 6

## HIA2\_DICDI

ID HIA2\_DICDI STANDARD; PRT; 117 AA.  
AC P42526;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Hisactophilin 2 (Histidine-rich actin-binding protein 2) (HS II).  
GN HATB.  
OS Dictyostelium discoideum (Slime mold).  
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.  
OX NCBI\_TaxID=44689;  
RN [1]  
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND POST-TRANSLATIONAL  
RP MODIFICATIONS.  
RC STRAIN=AX2, and AX3;  
RX MEDLINE=95122497; PubMed=7822284;  
RA Hanakam F., Eckerskorn C., Lottspeich F., Mueller-Taubenberger A.,  
RA Schaefer W., Gerisch G.;  
RT "The pH-sensitive actin-binding protein hisactophilin of  
RT Dictyostelium exists in two isoforms which both are myristoylated and  
RT distributed between plasma membrane and cytoplasm.";  
RL J. Biol. Chem. 270:596-602(1995).  
CC -!- FUNCTION: May act as an intracellular pH sensor that links  
CC chemotactic signals to responses in the microfilament system of  
CC the cells by nucleating actin polymerization or stabilizing the  
CC filaments.  
CC -!- SUBUNIT: Homodimer or heterodimer of hatA and hatB, linked by a  
CC disulfide bond.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic or associated with inner surface  
CC of plasma membrane.  
CC -!- PTM: Phosphorylated.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U13671; AAA66208.1; -.  
DR HSSP; P13231; 1HCE.  
DR DictyBase; DDB0001955; hatB.  
DR InterPro; IPR008999; Actin\_crosslink.

KW Actin-binding; Repeat; Myristate; Multigene family; Phosphorylation;  
KW Lipoprotein.

FT INIT\_MET 0 0  
FT LIPID 1 1 N-myristoyl glycine.  
FT DOMAIN 7 108 CONTAINS SEVERAL HHXH REPEATS.  
FT DOMAIN 33 85 2 X 13 AA APPROXIMATE REPEATS.  
FT REPEAT 33 45 1.  
FT REPEAT 73 85 2.  
SQ SEQUENCE 117 AA; 13503 MW; C92D50147FB8E0F8 CRC64;

Query Match 90.5%; Score 38; DB 1; Length 117;  
Best Local Similarity 100.0%; Pred. No. 6.1;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5  
|||  
Db 86 GHHHH 90

#### RESULT 7

##### INL3\_PIG

ID INL3\_PIG STANDARD; PRT; 131 AA.

AC P51461;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Insulin-like 3 precursor (Leydig insulin-like peptide) (Ley-I-L)

DE (Relaxin-like factor).

GN INSL3 OR RLF.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI\_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Testis;

RX MEDLINE=94075362; PubMed=8253799;

RA Adham I.M., Burkhardt E., Benahmed M., Engel W.;

RT "Cloning of a cDNA for a novel insulin-like peptide of the testicular

RT Leydig cells.";

RL J. Biol. Chem. 268:26668-26672(1993).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=94292172; PubMed=8020942;

RA Burkhardt E., Adham I.M., Brosig B., Gastmann A., Mattei M.-G.,

RA Engel W.;

RT "Structural organization of the porcine and human genes coding for a

RT Leydig cell-specific insulin-like peptide (LEY I-L) and chromosomal

RT localization of the human gene (INSL3).";

RL Genomics 20:13-19(1994).

CC -!- FUNCTION: Seems to play a role in testicular function. May be a

CC trophic hormone with a role in testicular descent in fetal life.

CC Is a ligand for LGR8 receptor (By similarity).

CC -!- SUBUNIT: Heterodimer of a B chain and an A chain linked by two

CC disulfide bonds (By similarity).

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Expressed exclusively in prenatal and



CC postnatal Leydig cells.  
 CC -!- SIMILARITY: Belongs to the insulin family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X73636; CAA52016.1; -.  
 DR EMBL; X68369; CAA48449.1; -.  
 DR PIR; A53024; A53024.  
 DR InterPro; IPR004825; Ins/IGF/relax.  
 DR Pfam; PF00049; Insulin; 1.  
 DR SMART; SM00078; IIGF; 1.  
 DR PROSITE; PS00262; INSULIN; 1.  
 KW Insulin family; Hormone; Signal.  
 FT SIGNAL 1 24 POTENTIAL.  
 FT CHAIN 25 56 INSULIN-LIKE 3 B CHAIN.  
 FT PROPEP 58 103 C PEPTIDE (POTENTIAL).  
 FT CHAIN 106 131 INSULIN-LIKE 3 A CHAIN.  
 FT DISULFID 34 116 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 46 129 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 115 120 BY SIMILARITY.  
 SQ SEQUENCE 131 AA; 14134 MW; 8AB718870859EF3A CRC64;

Query Match 90.5%; Score 38; DB 1; Length 131;  
 Best Local Similarity 100.0%; Pred. No. 6.9;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5  
 |||||  
 Db 99 GHHHH 103

# RESULT 8

## UREE\_PROMI

ID UREE\_PROMI STANDARD; PRT; 161 AA.  
 AC P17090;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Urease accessory protein ureE.  
 GN UREE.  
 OS Proteus mirabilis.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Proteus.  
 OX NCBI\_TaxID=584;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HI4320;  
 RX MEDLINE=90078080; PubMed=2687233;  
 RA Jones B.D., Mobley H.L.T.;  
 RT "Proteus mirabilis urease: nucleotide sequence determination and  
 RT comparison with jack bean urease.";

RL J. Bacteriol. 171:6414-6422(1989).  
 CC -!- FUNCTION: INVOLVED IN UREASE METALLOCENTER ASSEMBLY. BINDS NICKEL.  
 CC PROBABLY FUNCTIONS AS A NICKEL DONOR DURING METALLOCENTER ASSEMBLY  
 CC (BY SIMILARITY).  
 CC -!- SUBUNIT: Homodimer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -!- SIMILARITY: Belongs to the ureE family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M31834; AAA25670.1; -.  
 DR PIR; E43719; E43719.  
 DR InterPro; IPR007864; UreE\_C.  
 DR InterPro; IPR004029; UreE\_N.  
 DR Pfam; PF05194; UreE\_C; 1.  
 DR Pfam; PF02814; UreE\_N; 1.  
 KW Chaperone; Nickel.  
 FT DOMAIN 153 161 HIS-RICH CLUSTER EXPECTED TO FUNCTION AS  
 FT A NICKEL-BINDING SITE.  
 SQ SEQUENCE 161 AA; 17887 MW; 0126E60CF1B22BBF CRC64;

Query Match 90.5%; Score 38; DB 1; Length 161;  
 Best Local Similarity 100.0%; Pred. No. 8.5;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5  
 |||||  
 Db 152 GHHHH 156

# RESULT 9

## YOHM\_ECOLI

ID YOHM\_ECOLI STANDARD; PRT; 274 AA.  
 AC P76425; O08015;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein yohM.  
 GN YOHM OR B2106.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;

RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE=97251358; PubMed=9097040;  
 RA Itoh T., Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,  
 RA Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,  
 RA Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,  
 RA Nashimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y.,  
 RA Sivasundaram S., Tagami H., Takeda J., Takemoto K., Wada C.,  
 RA Yamamoto Y., Horiuchi T.;  
 RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome  
 RT corresponding to the 40.1-50.0 min region on the linkage map.";  
 RL DNA Res. 3:379-392(1996).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -!- SIMILARITY: TO M.JANNASCHII MJ1092 AND SOME, TO H.INFLUENZAE  
 CC HI1248.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AE000299; AAC75167.1; -.  
 DR EMBL; D90848; BAA15973.1; -.  
 DR PIR; A64978; A64978.  
 DR EcoGene; EG14071; yohM.  
 DR InterPro; IPR004688; NicO.  
 DR Pfam; PF03824; NicO; 1.  
 KW Hypothetical protein; Transmembrane; Complete proteome.  
 FT TRANSMEM 13 33 POTENTIAL.  
 FT TRANSMEM 57 77 POTENTIAL.  
 FT TRANSMEM 87 107 POTENTIAL.  
 FT TRANSMEM 176 196 POTENTIAL.  
 FT TRANSMEM 210 230 POTENTIAL.  
 FT TRANSMEM 252 272 POTENTIAL.  
 FT DOMAIN 127 146 POLY-HIS.  
 SQ SEQUENCE 274 AA; 30419 MW; 82C99F416E254C59 CRC64;

Query Match 90.5%; Score 38; DB 1; Length 274;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5  
 |||||  
 Db 139 GHHHH 143

# RESULT 10

MOX2\_XENLA

ID MOX2\_XENLA STANDARD; PRT; 298 AA.

AC P39021;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Homeobox protein MOX-2.  
 GN MOX2.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipioidea; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94232829; PubMed=7909944;  
 RA Candia A.F., Kovalik J.-P., Wright C.V.E.;  
 RT "Amino acid sequence of Mox-2 and comparison to its Xenopus and rat  
 RT homologs.";  
 RL Nucleic Acids Res. 21:4982-4982(1993).  
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).  
 CC -!- SIMILARITY: Contains 1 homeobox domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L20432; AAB00146.1; -.  
 DR HSSP; P14653; 1B72.  
 DR TRANSFAC; T04053; -.  
 DR InterPro; IPR001356; Homeobox.  
 DR InterPro; IPR000047; HTH\_lambprepressr.  
 DR Pfam; PF00046; homeobox; 1.  
 DR PRINTS; PR00024; HOMEBOX.  
 DR PRINTS; PR00031; HTHREPRESSR.  
 DR ProDom; PD000010; Homeobox; 1.  
 DR SMART; SM00389; HOX; 1.  
 DR PROSITE; PS00027; HOMEBOX\_1; 1.  
 DR PROSITE; PS50071; HOMEBOX\_2; 1.  
 KW Homeobox; DNA-binding; Nuclear protein; Developmental protein.  
 FT DOMAIN 42 47 POLY-SER.  
 FT DOMAIN 63 82 GLN/HIS-RICH (OPA-REPEAT).  
 FT DOMAIN 68 76 POLY-HIS.  
 FT DOMAIN 77 82 POLY-GLN.  
 FT DNA\_BIND 181 240 HOMEBOX.  
 SQ SEQUENCE 298 AA; 33245 MW; 154123DDED90824F CRC64;

Query Match 90.5%; Score 38; DB 1; Length 298;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5  
 |||||  
 Db 67 GHHHH 71

RESULT 11

HYPB\_RHILV  
 ID HYPB\_RHILV STANDARD; PRT; 299 AA.  
 AC P28155;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hydrogenase nickel incorporation protein hypB.  
 GN HYPB OR HUPM.  
 OS Rhizobium leguminosarum (biovar viciae).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.  
 OX NCBI\_TaxID=387;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=128c53;  
 RX MEDLINE=93316844; PubMed=8326860;  
 RA Rey L., Murillo J., Hernando Y., Hidalgo E., Cabrera E., Imperial J.,  
 RA Ruiz-Argueso T.;  
 RT "Molecular analysis of a microaerobically induced operon required for  
 RT hydrogenase synthesis in Rhizobium leguminosarum biovar viciae.";  
 RL Mol. Microbiol. 8:471-481(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=B10;  
 RA Brito B., Palacios J.M., Imperial J., Ruiz-Argueso T., Yang W.C.,  
 RA Bisseling T., Schmitt H., Kerl V., Bauer T., Kokotek W., Lotz W.;  
 RT "Organization of the hup-region and its differential transcription  
 RT in non-symbiotic and symbiotic cells of Rhizobium leguminosarum  
 RT bv. viciae B10.";  
 RL Mol. Plant Microbe Interact. 8:235-240(1997).  
 CC -!- FUNCTION: COULD BE INVOLVED IN NICKEL BINDING AND ACCUMULATION.  
 CC BINDS 3.9 NICKEL IONS PER MOLECULE.  
 CC -!- SIMILARITY: Belongs to the hypB/hupM family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X52974; CAA37160.1; -.  
 DR EMBL; Z36981; CAA85442.1; -.  
 DR PIR; S32874; S32874.  
 DR InterPro; IPR004392; HypB.  
 DR InterPro; IPR002894; HypB\_UreG.  
 DR Pfam; PF01495; HypB\_UreG; 1.  
 DR TIGRFAMS; TIGR00073; hypB; 1.  
 KW Metal-binding; Nickel.  
 FT DOMAIN 15 53 HIS-RICH.  
 SQ SEQUENCE 299 AA; 32590 MW; 5B7A53059D92E87D CRC64;

Query Match 90.5%; Score 38; DB 1; Length 299;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy            1 GHHHH 5  
               |||||  
 Db           25 GHHHH 29

RESULT 12

HYPB\_BRAJA

ID HYPB\_BRAJA        STANDARD;        PRT;    302 AA.

AC Q45257;

DT 01-NOV-1997 (Rel. 35, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Hydrogenase nickel incorporation protein hypB.

GN HYPB OR BLL6931.

OS Bradyrhizobium japonicum.

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;

OC Bradyrhizobiaceae; Bradyrhizobium.

OX NCBI\_TaxID=375;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=USDA 110;

RX MEDLINE=94137733; PubMed=8305450;

RA Fu C., Maier R.J.;

RT "Nucleotide sequences of two hydrogenase-related genes (hypA and

RT hypB) from Bradyrhizobium japonicum, one of which (hypB) encodes an

RT extremely histidine-rich region and guanine nucleotide-binding

RT domains.";

RL Biochim. Biophys. Acta 1184:135-138(1994).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=USDA 110;

RX MEDLINE=22484998; PubMed=12597275;

RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,

RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,

RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,

RA Tabata S.;

RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium

RT Bradyrhizobium japonicum USDA110.";

RL DNA Res. 9:189-197(2002).

CC -!- FUNCTION: May work in the mobilization of nickel into hydrogenase

CC enzyme. Binds 9 nickel ions per molecule.

CC -!- SIMILARITY: Belongs to the hypB/hupM family.

CC

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; L24513; AAA17763.1; -.

DR EMBL; AP005960; BAC52196.1; -.

DR InterPro; IPR004392; HypB.

DR InterPro; IPR002894; HypB\_UreG.

DR Pfam; PF01495; HypB\_UreG; 1.

DR TIGRFAMs; TIGR00073; hypB; 1.

KW Metal-binding; Nickel; Complete proteome.  
 FT DOMAIN 16 54 HIS-RICH.  
 FT CONFLICT 72 72 A -> T (IN REF. 1).  
 SQ SEQUENCE 302 AA; 32708 MW; D3B5F54F24AB90AA CRC64;

Query Match 90.5%; Score 38; DB 1; Length 302;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GH HHH 5  
 |||||  
 Db 34 GH HHH 38

# RESULT 13

## MOX2\_HUMAN

ID MOX2\_HUMAN STANDARD; PRT; 303 AA.

AC P50222; Q9UPL6;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Homeobox protein MOX-2 (Mesenchyme homeobox 2) (Growth arrest-specific  
 DE homeobox).

GN MEOX2 OR MOX2 OR GAX.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Embryo;

RX MEDLINE=95331791; PubMed=7607679;

RA Grigoriou M., Kastrinaki M.-C., Modi W., Theodorakis K., Mankoo B.,

RA Pachnis V., Karagogeos D.;

RT "Isolation of the human MOX2 homeobox gene and localization to  
 RT chromosome 7p22.1-p21.3.";

RL Genomics 26:550-555(1995).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Heart;

RX MEDLINE=95229154; PubMed=7713505;

RA Lepage D.F., Walsh K.;

RT "Molecular cloning and localization of the human GAX gene to 7p21.";

RL Genomics 24:535-540(1994).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Skin;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,

RA Bošak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP SEQUENCE OF 230-303 FROM N.A.  
 RA Cordes M., Lacy M.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Role in mesoderm induction and its earliest regional  
 CC specification, somitogenesis, and myogenic and sclerotomal  
 CC differentiation. May have a regulatory role when quiescent  
 CC vascular smooth muscle cells reenter the cell cycle (By  
 CC similarity).  
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).  
 CC -!- TISSUE SPECIFICITY: Embryo and placenta.  
 CC -!- SIMILARITY: Contains 1 homeobox domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X82629; CAA57949.1; -.  
 DR EMBL; L36328; AAA58497.1; -.  
 DR EMBL; BC017021; AAH17021.1; -.  
 DR EMBL; AC004452; AAC06184.1; -.  
 DR PIR; A55641; A55641.  
 DR PIR; A56837; A56837.  
 DR HSSP; P14653; 1B72.  
 DR TRANSFAC; T04005; -.  
 DR Genew; HGNC:7014; MEOX2.  
 DR MIM; 600535; -.  
 DR GO; GO:0008015; P:circulation; TAS.  
 DR GO; GO:0007275; P:development; TAS.  
 DR InterPro; IPR001356; Homeobox.  
 DR InterPro; IPR000047; HTH\_lambrepresr.  
 DR Pfam; PF00046; homeobox; 1.  
 DR PRINTS; PR00024; HOMEBOX.  
 DR PRINTS; PR00031; HTHREPRESSR.  
 DR ProDom; PD000010; Homeobox; 1.  
 DR SMART; SM00389; HOX; 1.  
 DR PROSITE; PS00027; HOMEBOX\_1; 1.  
 DR PROSITE; PS50071; HOMEBOX\_2; 1.  
 KW Homeobox; DNA-binding; Nuclear protein; Developmental protein.  
 FT DOMAIN 42 47 POLY-SER.  
 FT DOMAIN 68 79 POLY-HIS.  
 FT DOMAIN 80 85 POLY-GLN.



FT DNA\_BIND 186 245 HOMEBOX.  
 FT CONFLICT 58 58 G -> D (IN REF. 2).  
 FT CONFLICT 79 79 MISSING (IN REF. 2).  
 SQ SEQUENCE 303 AA; 33457 MW; 809ADE0CD090023D CRC64;

Query Match 90.5%; Score 38; DB 1; Length 303;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5  
 |||||  
 Db 67 GHHHH 71

# RESULT 14

## MOX2\_MOUSE

ID MOX2\_MOUSE STANDARD; PRT; 303 AA.

AC P32443;

DT 01-OCT-1993 (Rel. 27, Created)

DT 01-OCT-1993 (Rel. 27, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Homeobox protein MOX-2 (Mesenchyme homeobox 2).

GN MEOX2 OR MOX2 OR MOX-2 OR GAX.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=93201999; PubMed=1363541;

RA Candia A.F., Hu J., Crosby J., Lalley P.A., Noden D., Nadeau J.H.,

RA Wright C.V.E.;

RT "Mox-1 and Mox-2 define a novel homeobox gene subfamily and are

RT differentially expressed during early mesodermal patterning in mouse

RT embryos.";

RL Development 116:1123-1136(1992).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=94232829; PubMed=7909944;

RA Candia A.F., Kovalik J.-P., Wright C.V.E.;

RT "Amino acid sequence of Mox-2 and comparison to its Xenopus and rat

RT homologs.";

RL Nucleic Acids Res. 21:4982-4982(1993).

RN [3]

RP SEQUENCE OF 1-11 FROM N.A.

RX MEDLINE=95349593; PubMed=7623821;

RA Andres V., Fisher S., Wearsch P., Walsh K.;

RT "Regulation of Gax homeobox gene transcription by a combination of

RT positive factors including myocyte-specific enhancer factor 2.";

RL Mol. Cell. Biol. 15:4272-4281(1995).

CC -!- FUNCTION: Role in mesoderm induction and its earliest regional

CC specification, somitogenesis, and myogenic and sclerotomal

CC differentiation. May have a regulatory role when quiescent

CC vascular smooth muscle cells reenter the cell cycle.

CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).

CC -!- DEVELOPMENTAL STAGE: It is not expressed before 8-8.5 dpc. At 8-

CC 8.5 dpc it is found on the entire epithelium of the somite. At 9.5

CC dpc its expression is restricted to the sclerotome. At 10.5 dpc it  
 CC is found in sclerotomally derived cells including the vertebral  
 CC and costal precursors.  
 CC -!- SIMILARITY: Contains 1 homeobox domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z16406; CAA78899.1; -.  
 DR EMBL; S79168; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; B49122; B49122.  
 DR HSSP; P14653; 1B72.  
 DR TRANSFAC; T04048; -.  
 DR MGD; MGI:103219; Meox2.  
 DR InterPro; IPR001356; Homeobox.  
 DR InterPro; IPR000047; HTH\_lambprepressr.  
 DR Pfam; PF00046; homeobox; 1.  
 DR PRINTS; PR00024; HOMEBOX.  
 DR PRINTS; PR00031; HTHREPRESSR.  
 DR ProDom; PD000010; Homeobox; 1.  
 DR SMART; SM00389; HOX; 1.  
 DR PROSITE; PS00027; HOMEBOX\_1; 1.  
 DR PROSITE; PS50071; HOMEBOX\_2; 1.  
 KW Homeobox; DNA-binding; Nuclear protein; Developmental protein.  
 FT DOMAIN 42 47 POLY-SER.  
 FT DOMAIN 68 79 POLY-HIS.  
 FT DOMAIN 80 85 POLY-GLN.  
 FT DOMAIN 63 85 GLN/HIS-RICH (OPA-REPEAT).  
 FT DNA\_BIND 186 245 HOMEBOX.  
 SQ SEQUENCE 303 AA; 33506 MW; 41BD05FC39AA4427 CRC64;

Query Match 90.5%; Score 38; DB 1; Length 303;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GHHHH 5  
 |||||  
 Db 67 GHHHH 71

#### RESULT 15

##### MOX2\_RAT

ID MOX2\_RAT STANDARD; PRT; 303 AA.  
 AC P39020;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Homeobox protein MOX-2 (Growth arrest-specific homeobox).  
 GN MEOX2 OR MOX2 OR MOX-2 OR GAX.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Aorta;  
 RX MEDLINE=93268321; PubMed=8098844;  
 RA Gorski D.H., Lepage D.F., Patel C.V., Copeland N.G., Jenkins N.A.,  
 RA Walsh K.;  
 RT "Molecular cloning of a diverged homeobox gene that is rapidly down-  
 RT regulated during the G0/G1 transition in vascular smooth muscle  
 RT cells.";  
 RL Mol. Cell. Biol. 13:3722-3733(1993).  
 RN [2]  
 RP REVISIONS.  
 RA Walsh K.;  
 RL Submitted (MAR-1993) to the EMBL/GenBank/DDBJ databases.  
 CC !- FUNCTION: Role in mesoderm induction and its earliest regional  
 CC specification, somitogenesis, and myogenic and sclerotomal  
 CC differentiation. May have a regulatory role when quiescent  
 CC vascular smooth muscle cells reenter the cell cycle.  
 CC !- SUBCELLULAR LOCATION: Nuclear (Potential).  
 CC !- TISSUE SPECIFICITY: Aorta and heart. Also detected in lung and  
 CC kidney.  
 CC !- INDUCTION: Rapidly and transiently down-regulated during the  
 CC transition from G0 to G1 induced by mitogen stimulation.  
 CC !- SIMILARITY: Contains 1 homeobox domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z17223; CAA78931.1; -.  
 DR PIR; A48130; A48130.  
 DR HSSP; P14653; 1B72.  
 DR TRANSFAC; T04054; -.  
 DR InterPro; IPR001356; Homeobox.  
 DR InterPro; IPR000047; HTH\_lambrepressr.  
 DR Pfam; PF00046; homeobox; 1.  
 DR PRINTS; PR00024; HOMEBOX.  
 DR PRINTS; PR00031; HTHREPRESSR.  
 DR ProDom; PD000010; Homeobox; 1.  
 DR SMART; SM00389; HOX; 1.  
 DR PROSITE; PS00027; HOMEBOX\_1; 1.  
 DR PROSITE; PS50071; HOMEBOX\_2; 1.  
 KW Homeobox; DNA-binding; Nuclear protein; Developmental protein.  
 FT DOMAIN 42 47 POLY-SER.  
 FT DOMAIN 68 79 POLY-HIS.  
 FT DOMAIN 80 85 POLY-GLN.  
 FT DOMAIN 64 85 GLN/HIS-RICH (OPA-REPEAT).  
 FT DNA\_BIND 186 245 HOMEBOX.  
 SQ SEQUENCE 303 AA; 33605 MW; 7776642AEFA3A2E8 CRC64;

Query Match 90.5%; Score 38; DB 1; Length 303;  
 Best Local Similarity 100.0%; Pred. No. 16;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 GHHHH 5
        |||||
Db      67 GHHHH 71
```

Search completed: March 5, 2004, 16:23:43  
Job time : 1.81481 secs